

BLOCK 1	
INTRODUCTION TO MANAGERIAL ECONOMICS	5
BLOCK 2	
DEMAND AND REVENUE ANALYSIS	69
BLOCK 3	
PRODUCTION AND COST ANALYSIS	131
BLOCK 4	
PRICING DECISIONS	223

COURSE DESIGN AND PREPARATION TEAM

Prof. K. Ravi Sankar
Director, SOMS,
IGNOU, New Delhi

Pro. Qamar Ahsan
Retd. Prof. of Economics and Vice Chancellor,
Magadh University Bihar

Prof. K.V. Bhanu Murthy
Former Dean
Faculty of Commerce and Business
Delhi School of Economics, Delhi

Prof. Sunitha Raju
Indian Institute of Foreign Trade
New Delhi

Dr. Utpal Chattopadhyay
Associate Professor
National Institute of Industrial Engineering
Mumbai

Prof. Biswajit Nag
Indian Institute of Foreign Trade
New Delhi

Dr. Alka Mittal
Associate Professor
Guru Gobind Singh Indraprastha University
Delhi

Prof. G. Subbayamma
School of Management Studies,
IGNOU, New Delhi

Prof. Srilatha
School of Management Studies,
IGNOU, New Delhi

Prof. Neeti Agrawal
School of Management Studies,
IGNOU, New Delhi

Prof. Anjali C. Ramteke
School of Management Studies
IGNOU, New Delhi

*Prof. V. L. Mote (Retd.)
IIM, Ahmedabad

*Prof. G.S. Gupta
IIM, Ahmedabad

*Prof. Arindam Banik
IMI, New Delhi

*Prof. Atmanand
MDI, Gurgaon

*Prof. Rajat Kathuria
IMI, New Delhi

*Dr. C.G. Naidu
Planning & Development Division
IGNOU, New Delhi

Prof. Kamal Vagrecha
School of Management Studies,
IGNOU, New Delhi

Prof. Nayantara Padhi
School of Management Studies
IGNOU, New Delhi

Sh. T.V. Vijay Kumar
School of Management Studies
IGNOU, New Delhi

Course Coordinator and Editor:

Dr. Leena Singh
SOMS, IGNOU
New Delhi

Acknowledgement : Parts of this course is adapted from the earlier MS-09: Managerial Economics course and the persons marked with (*) were the original contributors and the profiles are as it was in that version.

PRINT PRODUCTION

Mr. Tilak Raj,
Assistant Registrar
MPDD, IGNOU, New Delhi

February, 2022

© Indira Gandhi National Open University, 2022

ISBN :

All rights reserved. No part of this work may be reproduced in any form, by mimeograph or any other means, without permission in writing from the Indira Gandhi National Open University. Further information on the Indira Gandhi National Open University courses may be obtained from the University's office at Maidan Garhi, New Delhi – 110 068

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi, by the Registrar, MPDD, IGNOU.

Laser Typesetting : Akashdeep Printers, 20-Ansari Road, Daryaganj, New Delhi-110002

Printed at :

MMPC-010: MANAGERIAL ECONOMICS

Managerial Economics can be viewed as an application of that part of microeconomics that focuses on topics such as risk, demand, production, cost, pricing, and market structure. Understanding these principles will help to develop a rational decision-making perspective and will sharpen the analytical framework that the executive must bring to bear on managerial decisions.

Individuals and firms interact in both the product and the factor markets. Prices of outputs and inputs are determined in these markets and guide the decisions of all market participants. The firm is an entity that organizes factors of production in order to produce goods and services to meet the demands of consumers and other firms. In a market system, the interplay of individuals and firms is not subject to central control. The prices of both products and factors of production guide this interaction. Within firms, however, transactions and information costs are reduced. The size of the firm is limited because transaction costs within the firm will rise as the firm grows, and because management skill is limited.

It is assumed that the goal of the firm is to maximize the value of the firm or the present value of all future profits, defined as revenue less all costs, explicit and implicit. Implicit costs such as the remuneration and interest that owners and managers earn are not accounted and may result in an inefficient allocation of resources. The objective of profit maximization is subject to legal, moral, contractual, financial, and technological constraints. Some economists argue that the firm's objective is a "satisfactory" level of profit rather than maximum profit. The principal-agent problem arises where the owner of a firm and the manager of that firm have different objectives. The problem can be solved by tying part of the manager's salary to profits and /or changes in the price of the firm's stock.

Profit plays two primary roles in the free-market system. First, it acts as a signal to producers to increase or decrease the rate of output, or to enter or leave an industry. Second, profit is a reward for entrepreneurial activity, including risk taking and innovation. In a competitive industry, economic profits tend to be transitory. The achievement of high profits by a firm usually results in other firms increasing their output of that product, thus reducing price and profit. Firms that have monopoly power may be able to earn above-normal profits over a longer period; such profit does not play a socially useful role in the economy. Different types of pricing strategies are adopted, keeping in mind, the varying market structure for products.

The primary role of economics in management is in making optimizing decisions where constraints apply. The application of the principles of managerial economics will help manager ensure that resources are allocated efficiently within the firm, and that the firm reacts to changes in the economic environment.

Block 1

INTRODUCTION TO MANAGERIAL ECONOMICS

Unit 1: Scope of Managerial Economics	9
Unit 2: The Firm: Stakeholders, Objectives and Decisions Issues	22
Unit 3: Basic Concepts and Techniques	42

BLOCK 1 INTRODUCTION TO MANAGERIAL ECONOMICS

This block provided an introduction to managerial economics and also highlights its importance in the current economic environment. With the easing of controls on the economy and with increasing amount of output being generated through the market, firms have begun to exert positive influence in the production of goods and services in India. The change began in 1991 with economic reforms and has now acquired a momentum of its own. It is in the context of growing liberalization of the Indian economy that a student of managerial economics will need to understand and appreciate the importance of the subject. **Block 1** provides the student with the basic tool kit that aids understanding of the nature and process of decision making by a manager.

Unit 1 defines the nature and scope of managerial economics. Economics is a vast subject with a number of branches that have developed over time. This includes economics of industry, public finance, international trade, and agricultural economics. Broadly, economics may be divided into macroeconomics and microeconomics. Macroeconomics, as the name suggests, is the study of the overall economy and its aggregates such as Gross National Product, Inflation, Unemployment, Exports, Imports, and Taxation Policy etc. On the other hand, microeconomics deals with individual actors in the economy such as firms and individuals. Managerial economics can be thought of an applied microeconomics.

Unit 2 introduces you to one of the major units of analysis of managerial economics, considering a firm. The objectives of a firm are defined and explained and a framework is provided for efficient allocation of resources within a firm. The firm occupies centre stage in a market oriented economy and understanding its rationale for existence and the nature of its operations will be explored in this unit. You should, however, appreciate that since managerial economics chiefly focuses on the firm, different aspects of firm's behavior will be found in other units of this course as well.

The concepts and techniques developed in **Unit 3** are designed to increase the effectiveness of decision making by providing an analytical framework used by managers. Thus emphasis is placed on optimization techniques, both constrained and unconstrained as well as an introduction to decision making under uncertainty.

UNIT 1 SCOPE OF MANAGERIAL ECONOMICS

Objectives

After studying this unit, you should be able to:

- **understand** the nature and scope of managerial economics;
- **familiarize** yourself with economic terminology;
- **develop** some insight into economic issues;
- **acquire** some information about economic institutions;
- **understand** the concept of trade-offs or policy options facing society today.

Structure

- 1.1 Introduction
- 1.2 Fundamental Nature of Managerial Economics
- 1.3 Scope of Managerial Economics
- 1.4 Appropriate Definitions
- 1.5 Managerial Economics and other Disciplines
- 1.6 Economic Analysis
- 1.7 Basic Characteristics: Decision-Making
- 1.8 Summary
- 1.9 Self-Assessment Questions
- 1.10 Further Readings

1.1 INTRODUCTION

For most purposes economics can be divided into two broad categories, microeconomics and macroeconomics. Macroeconomics as the name suggests is the study of the overall economy and its aggregates such as Gross National Product, Inflation, Unemployment, Exports, Imports, Taxation Policy etc. Macroeconomics addresses questions about changes in investment, government spending, employment, prices, exchange rate of the rupee and so on. Importantly, only aggregate levels of these variables are considered in the study of macroeconomics. But hidden in the aggregate data are changes in output of a number of individual firms, the consumption decision of consumers like you, and the changes in the prices of particular goods and services.

Although macroeconomic issues are important and occupy the time of media and command the attention of the newspapers, micro aspects of the economy are also important and often are of more direct application to the day to day problems facing a manager. Microeconomics deals with individual actors in the economy such as firms and individuals. Managerial economics can be thought of as applied microeconomics and its focus is on the interaction of firms and individuals in markets.

When you read a newspaper or switch on a television, you hear economic terminology used with increasing regularity. For a manager, some of these economic terms are of direct relevance and therefore it is essential to not only

understand them but also apply them in relevant situations. For example, GDP growth rate could impact the product a manager is marketing, change in money supply by the RBI could impact inflation and affect the demand for your product, fiscal deficit could affect interest rates and therefore investment spending by a manager etc. The focus of managerial economics is on how the firm reacts to changes in the economic environment in which it operates and how it predicts these changes and devises the best possible strategies to achieve the objectives that underlie its existence.

The economy is the institutional structure through which individuals and firms in a society coordinate their desires. Economics is the study of how human beings in a society go about achieving their wants and desires. It is also defined as the study of allocation of scarce resources to satisfy individual wants or desires. The latter is perhaps the best way to broadly define the study of economics in general. The emphasis is on allocation of scarce resources across competing ends. You should recognize that human wants are unlimited and therefore choice is necessary. Choices necessarily involve trade-offs. For example, if you wish to acquire an MBA degree, you must take time off to devote to study. Your time has many uses and when you devote more time to study you are allocating it to a particular use in order to achieve your goal. Economics would be a most uninteresting subject if resources were unlimited and no trade offs was involved in decision making.

There are many general insights economists have gained into how the economy functions. Economic theory ties together economists' terminology and knowledge about economic institutions. An economic institution is a physical or mental structure that significantly influences economic decisions. Corporations, governments, markets are all economic institutions. Similarly cultural norms are the standards people use when they determine whether a particular activity or behaviour is acceptable. For example, Hindus avoid meat and fish on Tuesdays. This has an economic dimension as it has a direct impact on the sale of these items on Tuesdays. Further, economic policy is the action usually taken by the government, to influence economic events. And finally, economic reasoning helps in thinking like an economist. Economists analyze questions and issues on the basis of trade-offs i.e. they compare the cost and the benefits of every issue and make decisions based on those costs and benefits.

The market is perhaps the single most important and complex institution in our economy. A market is not necessarily a physical location, but a description of any state that involves exchange. The exchange could be instantaneous or it could be over time i.e. exchange which is agreed today but where the transaction takes place, say after 3 months. You will learn in this course the myriad functions that markets perform, most significantly bringing buyers and sellers together. Markets could be competitive or monopolistic, with a large number of firms or a small number of firms, with free entry and exit or government licensing restricting entry of firms and so on. The major point is that firms operate in different types of markets and use the well-established principles of managerial economics to improve profitability. Managerial economics draws on economic analysis for such concepts as cost, demand, profit and competition. It attempts to bridge the gap between the purely analytical problems that intrigue many economic theorists and the day-to-day decisions that managers must face. It offers powerful tools and

approaches for managerial policy-making. It will be relevant to present here several examples illustrating the problems that managerial economics can help to address. These also explain how managerial economics is an integral part of business. Demand, supply, cost, production, market, competition, price etc. are important concepts in real business decisions.

1.2 FUNDAMENTAL NATURE OF MANAGERIAL ECONOMICS

A close relationship between management and economics has led to the development of managerial economics. Management is the guidance, leadership and control of the efforts of a group of people towards some common objective. While this description does inform about the purpose or function of management, it tells us little about the nature of the management process. Koontz and O'Donnell define management as the creation and maintenance of an internal environment in an enterprise where individuals, working together in groups, can perform efficiently and effectively towards the attainment of group goals. Thus, management is

- Coordination
- An activity or an ongoing process
- A purposive process
- An art of getting things done by other people.

On the other hand, economics as stated above is engaged in analyzing and providing answers to manifestations of the most fundamental problem of scarcity. Scarcity of resources results from two fundamental facts of life:

- Human wants are virtually unlimited and insatiable, and
- Economic resources to satisfy these human demands are limited.

Thus, we cannot have everything we want; we must make choices broadly in regard to the following:

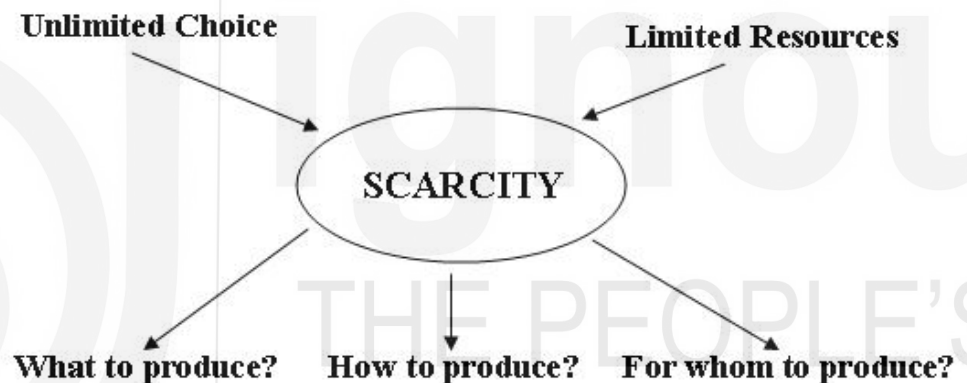
- What to produce?
- How to produce? and
- For whom to produce?

These three choice problems have become the three central issues of an economy as shown in figure 1.1. Economics has developed several concepts and analytical tools to deal with the question of allocation of scarce resources among competing ends. The non-trivial problem that needs to be addressed is how an economy through its various institutions solves or answers the three crucial questions posed above. There are three ways by which this can be achieved. One, entirely by the market mechanism, two, entirely by the government or finally, and more reasonably, by a combination of the first two approaches. Realistically all economies employ the last option, but the relative roles of the market and government vary across countries. For example, in India the market has started playing a more important role in the economy while the government has begun to withdraw from certain activities. Thus, the market mechanism is gaining importance. A similar

change is happening all over the world, including in China. But there are economies such as Myanmar and Cuba where the government still plays an overwhelming part in solving the resource allocation problem. Essentially, the market is supposed to guide resources to their most efficient use. For example if the salaries earned by MBA degree holders continue to rise, there will be more and more students wanting to earn the degree and more and more institutes wanting to provide such degrees to take advantage of this opportunity. The government may not force this to happen, it will happen on its own through the market mechanism. The government, if anything, could provide a regulatory function to ensure quality and consumer protection.

According to the central deduction of economic theory, under certain conditions, markets allocate resources efficiently. 'Efficiency' has a special meaning in this context. The theory says that markets will produce an outcome such that, given the economy's scarce resources, it is impossible to make anybody better-off without making somebody else worse-off. Scope of Managerial

Figure 1.1: Three Choice Problems of an Economy



In rich countries, markets are too familiar to attract attention. Yet, certain awe is appropriate. Let us take an incident where Soviet planners visited a vegetable market in London during the early days of perestroika, they were impressed to find no queues, shortages, or mountains of spoiled and unwanted vegetables. They took their hosts aside and said: "We understand, you have to say it's all done by supply and demand. But can't you tell us what's really going on? Where are your planners and what are their methods?"

The essence of the market mechanism is indeed captured by the supply-and demand diagram that you will become familiar with in Block 4. At the place where the curves intersect, a price is set such that demand equals supply. There, and only there, the benefit from consuming one more unit exactly matches the cost of producing it. If output were less, the benefit from consuming more would exceed the cost of producing it. If output were higher, the cost of producing the extra units would exceed the extra benefits. So the point where supply equals demand is "efficient".

However, the conditions for market efficiency are extremely demanding—far too demanding ever to be met in the real world. The theory requires "perfect competition": there must be many buyers and sellers; goods from competing suppliers must be indistinguishable; buyers and sellers must be fully

informed; and markets must be complete—that is, there must be markets not just for bread here and now, but for bread in any state of the world. (What is the price today for a loaf to be delivered in Timbuktu on the second Tuesday in December 2020 if it rains?)

In other words, market failure is pervasive. It comes in four main varieties:

Monopoly: By reducing his sales, a monopolist can drive up the price of his good. His sales will fall but his profits will rise. Consumption and production are less than the efficient amount, causing a deadweight loss in welfare.

Public goods: Some goods cannot be supplied by markets. If you refuse to pay for a new coat, the seller will refuse to supply you. If you refuse to pay for national defence, the “good” cannot easily be withheld. You might be tempted to let others pay. The same reasoning applies to other “non-excludable” goods such as law and order, clean air, and so on. Since private sellers cannot expect to recover the costs of producing such goods, they will fail to supply them.

Externalities: Making some goods causes pollution: the cost is borne by people with no say in deciding how much to produce. Consuming some goods (education, anti-lock brakes) spreads benefits beyond the buyer; again, this will be ignored when the market decides how much to produce. In the case of “good” externalities, markets will supply too little; in the case of “bads”, too much.

Information: In some ways a special kind of externality, this deserves to be mentioned separately because of the emphasis placed upon it in recent economic theory. To see why information matters, consider the market for used cars. A buyer, lacking reliable information, may see the price as providing clues about a car’s condition. This puts sellers in a quandary: if they cut prices, they may only convince people that their cars are rubbish.

The labour market, many economists believe, is another such ‘market for lemons’. This may help to explain why it is so difficult for the unemployed to price themselves into work.

When markets fail, there is a case for intervention. But two questions need to be answered first. How much does market failure matter in practice? And can governments put the failure right? Markets often correct their own failures. In other cases, an apparent failure does nobody any harm. In general, market failure matters less in practice than is often supposed.

Monopoly, for instance, may seem to preclude an efficient market. This is wrong. The mere fact of monopoly does not establish that any economic harm is being done. If a monopoly is protected from would-be competitors by high barriers to entry, it can raise its prices and earn excessive profits. If that happens, the monopoly is undeniably harmful. But if barriers to entry are low, lack of actual (as opposed to potential) competitors does not prove that the monopoly is damaging: the threat of competition may be enough to make it behave as though it were a competitive firm. Many economists would accept that an American multinational technology corporation for instance, is a near-monopolist in some parts of the personal-computer software business—

yet would argue that the firm is doing no harm to consumers because its markets remain highly contestable. Because of that persistent threat of competition, the company prices its products keenly. In this and in other ways it behaves as though it were a smaller firm in a competitive market.

Even on economic grounds (never mind other considerations), there is no tidy answer to the question of where the boundary between state i.e. governments and market should lie. Markets do fail because of monopoly, public goods, externalities, lack of information and for other reasons. But, more than critics allow, markets find ways to mitigate the harm—and that is a task at which governments have often been strikingly unsuccessful. All in all, a strong presumption in favour of markets seems wise. This is not because classical economic theory says so, but because experience seems to agree. And as stated above, the real world seems to be moving in the direction of placing more reliance on markets than on governments.

1.3 SCOPE OF MANAGERIAL ECONOMICS

From the point of view of a firm, managerial economics, may be defined as economics applied to “problems of choice” or alternatives and allocation of scarce resources by the firms. Thus managerial economics is the study of allocation of resources available to a firm or a unit of management among the activities of that unit. Managerial economics is concerned with the application of economic concepts and analysis to the problem of formulating rational managerial decisions. There are four groups of problem in both decisions-making and forward planning.

Resource Allocation: Scarce resources have to be used with utmost efficiency to get optimal results. These include production programming and problem of transportation etc. How does resource allocation take place within a firm? Naturally, a manager decides how to allocate resources to their respective uses within the firm, while as stated above, the resource allocation decision outside the firm is primarily done through the market. Thus, one important insight you can draw about the firm is that within it resources are guided by the manager in a manner that achieves the objectives of the firm. More will be said about this in Unit 2.

Inventory and queuing problem: Inventory problems involve decisions about holding of optimal levels of stocks of raw materials and finished goods over a period. These decisions are taken by considering demand and supply conditions. Queuing problems involve decisions about installation of additional machines or hiring of extra labour in order to balance the business lost by not undertaking these activities.

Pricing Problem: Fixing prices for the products of the firm is an important decision-making process. Pricing problems involve decisions regarding various methods of prices to be adopted.

Investment Problem: Forward planning involves investment problems. These are problems of allocating scarce resources over time. For example, investing in new plants, how much to invest, sources of funds, etc.

Study of managerial economics essentially involves the analysis of certain major subjects like:

- The business firm and its objectives
- Demand analysis, estimation and forecasting
- Production and Cost analysis
- Pricing theory and policies
- Profit analysis with special reference to break-even point
- Capital budgeting for investment decisions
- Competition.

Demand analysis and forecasting help a manager in the earliest stage in choosing the product and in planning output levels. A study of demand elasticity goes a long way in helping the firm to fix prices for its products. The theory of cost also forms an essential part of this subject. Estimation is necessary for making output variations with fixed plants or for the purpose of new investments in the same line of production or in a different venture. The firm works for profits and optimal or near maximum profits depend upon accurate price decisions. Theories regarding price determination under various market conditions enable the firm to solve the price fixation problems. Control of costs, proper pricing policies, break-even analysis, alternative profit policies are some of the important techniques in profit planning for the firm which has to work under conditions of uncertainty. Thus managerial economics tries to find out which course is likely to be the best for the firm under a given set of conditions.

1.4 APPROPRIATE DEFINITIONS

According to **McNair and Meriam**, “Managerial economics is the use of economic modes of thought to analyze business situations.” According to **Prof. Evan J Douglas**, ‘Managerial economics’ is concerned with the application of economic principles and methodologies to the decision making process within the firm or organisation under the conditions of uncertainty”. **Spencer and Siegelman** define it as “The integration of economic theory with business practices for the purpose of facilitating decision making and forward planning by management.” According to **Hailstones and Rothwel**, “Managerial economics is the application of economic theory and analysis to practice of business firms and other institutions.” A common thread runs through all these descriptions of managerial economics which is using a framework of analysis to arrive at informed decisions to maximize the firm’s objectives, often in an environment of uncertainty. It is important to recognize that decisions taken while employing a framework of analysis are likely to be more successful than decisions that are knee jerk or gut feel decisions.

Activity 1

- a) Development of managerial economics is the result of close interrelationship between management and economics. Discuss.

.....
.....

- b) Which statement is true of the basic economic problem?
- (i) The problem will exist as long as resources are limited and desires are unlimited.
 - (ii) The problem exists only in less developed countries.
 - (iii) The problem will disappear as production expands.
 - (iv) The advancement of technology will cause the problem to disappear.
- c) Why is decision making by any management truly economic in nature?

1.5 MANAGERIAL ECONOMICS AND OTHER DISCIPLINES

Managerial economics is linked with various other fields of study like—

Microeconomic Theory: As stated in the introduction, the roots of managerial economics spring from micro-economic theory. Price theory, demand concepts and theories of market structure are few elements of micro economics used by managerial economists. It has an applied bias as it applies economic theories in order to solve real world problems of enterprises.

Macroeconomic Theory: This field has little relevance for managerial economics but at least one part of it is incorporated in managerial economics i.e. national income forecasting. The latter could be an important aid to business condition analysis, which in turn could be a valuable input for forecasting the demand for specific product groups.

Operations Research: This field is used in managerial economics to find out the best of all possibilities. Linear programming is a great aid in decision making in business and industry as it can help in solving problems like determination of facilities on machine scheduling, distribution of commodities and optimum product mix etc.

Theory of Decision Making: Decision theory has been developed to deal with problems of choice or decision making under uncertainty, where the applicability of figures required for the utility calculus are not available. Economic theory is based on assumptions of a single goal whereas decision theory breaks new grounds by recognizing multiplicity of goals and persuasiveness of uncertainty in the real world of management.

Statistics: Statistics helps in empirical testing of theory. With its help, better decisions relating to demand and cost functions, production, sales or

distribution are taken. Managerial economics is heavily dependent on statistical methods.

Management Theory and Accounting: Maximization of profit has been regarded as a central concept in the theory of the firm in microeconomics. In recent years, organisation theorists have talked about “satisficing” instead of “maximizing” as an objective of the enterprise. Accounting data and statements constitute the language of business. In fact the link is so close that “managerial accounting” has developed as a separate and specialized field in itself.

1.6 ECONOMIC ANALYSIS

Economic activity is the constant effort to match ends to means because of scarcity of resources. The optimal economic activity is to maximize the attainment of ends, the means and their scarcities or to minimize the use of resources, given the ends and their priorities.

Decision making by management is truly economic in nature because it involves choices among a set of alternatives - alternative courses of action. The optimal decision making is an act of optimal economic choice, considering objectives and constraints. This justifies an evaluation of managerial decisions through concepts, precepts, tools and techniques of economic analysis of the following types:

Micro and Macro Analysis: In micro-analysis the problem of choice is focused on single individual entities like a consumer, a producer, a market etc. Macro analysis deals with the problem in totality like national income, general price level etc.

Partial and General Equilibrium Analysis: To attain the state of stable equilibrium, the economic problem may be analyzed part by part - one at a time - assuming “other things remaining the same.” This is partial equilibrium analysis. In general equilibrium analysis the assumption of “given” or “other things remaining equal” may be relaxed and interdependence or interactions among variables may be allowed.

Static, Comparative Static and Dynamic Analysis: This is in reference to time dimension. A problem may be analyzed

- allowing no change at a point of time (static)
- allowing once for all change at a point of time (comparative static)
- allowing successive changes over a period of time (dynamic).

Positive and Normative Analysis: In positive economic analysis, the problem is analyzed in objective terms based on principles and theories. In normative economic analysis, the problem is analyzed based on value judgement (norms). In simple terms, positive analysis is ‘what it is’ and normative analysis is ‘what it should be.’ For example, CEOs in private Indian enterprises earn 15 times as much as the lowest paid employee is a positive statement, a description of what is. A normative statement would be that CEOs should be paid 4-5 times the lowest paid employee.

Activity 2

a) The major groups of problems in decision making are:

- (i)
- (ii)
- (iii)
- (iv)

b) The 3 choice problems of an economy are:

- (i)
- (ii)
- (iii)

The problems arise due to.....

c) Name the kind of economic analysis that is appropriate for each of the following:

- (i) The Furucture company has expansion plans
- (ii) The telecommunication company is making loss
- (iii) The real estate is facing recession
- (iv) The population growth in India is alarming
- (v) There is a bearish trend in the stock market

1.7 BASIC CHARACTERISTICS: DECISION-MAKING

Managerial Economics serves as ‘a link between traditional economics and the decision making sciences’ for business decision making.

The best way to get acquainted with managerial economics and decision making is to come face to face with real world decision problems.

Case Study of an American multinational retail company

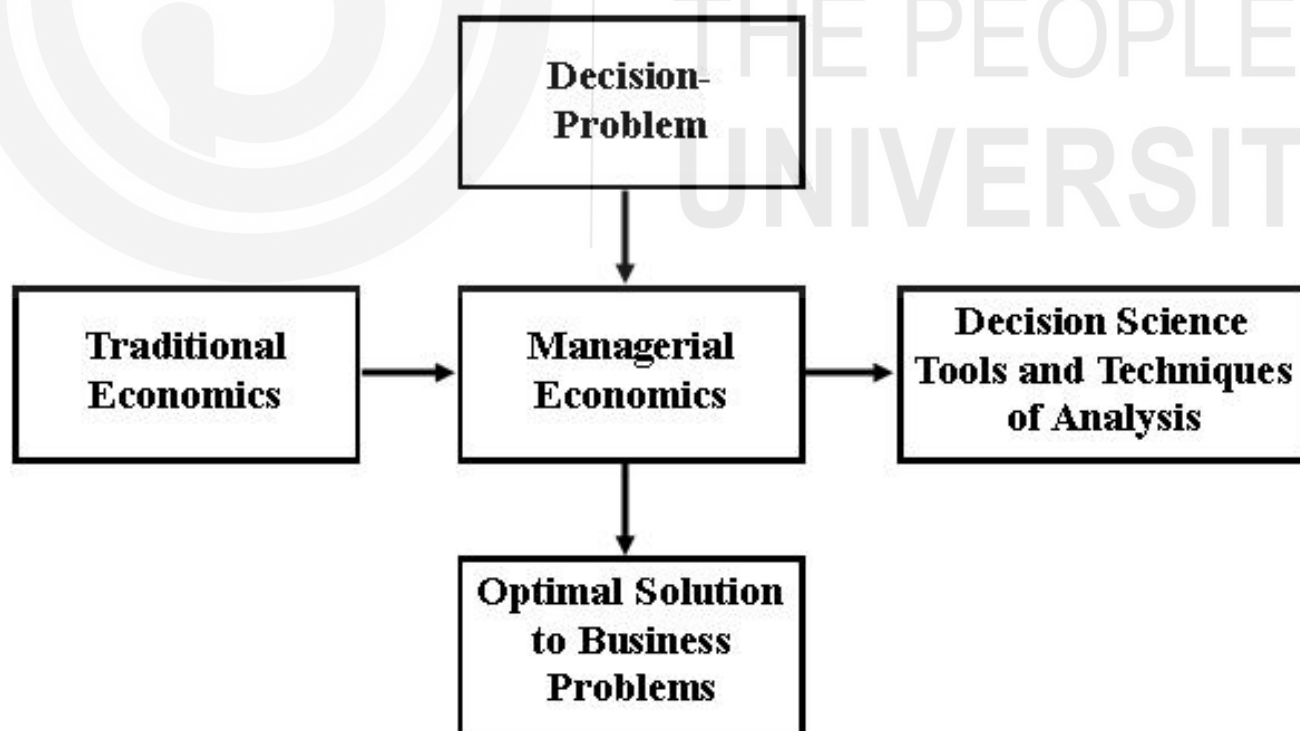
India is one of the biggest retail markets in the world in terms of size and the growth potential in future. An American multinational retail company looked at this potential and decided to choose among different alternatives of entering foreign markets such as strategic partnership, joint venture, FDI, wholly owned subsidiary etc. Joint ventures were chosen as the first step of entering the new market keeping in mind several economic and regulatory restrictions (of India as 100% FDI was not allowed in multi brand retailing). It decided to enter India in 2007 via joint venture with an Indian enterprise by setting up cash-and-carry wholesale outlets where it planned to supply to an Indian retail company which would help it in getting share in consumer market of India. But this did not fulfill the dream of the multinational and this venture did not survive for long and ended in 2014.

The American retail giant had two options; to leave India or stay back. The company stayed back as it looked at the consumer market potential and it had invested too much already. It kept looking at different options and then in 2018 it acquired 77% stake of an Indian e-commerce company for \$16 billion (about ₹1.05 lakh crore). The decision to acquire the Indian e-commerce company was well thought as it is one of the biggest e-commerce company in India which would give access to a large number of consumers of India and on the other hand would help in giving direct competition to its old rival an American e-commerce company. These types of major decision-making problems are faced by managers in business world and the managers have to choose the best alternative from the available alternatives for the firm.

The basic characteristics of managerial economics can now be enumerated as:

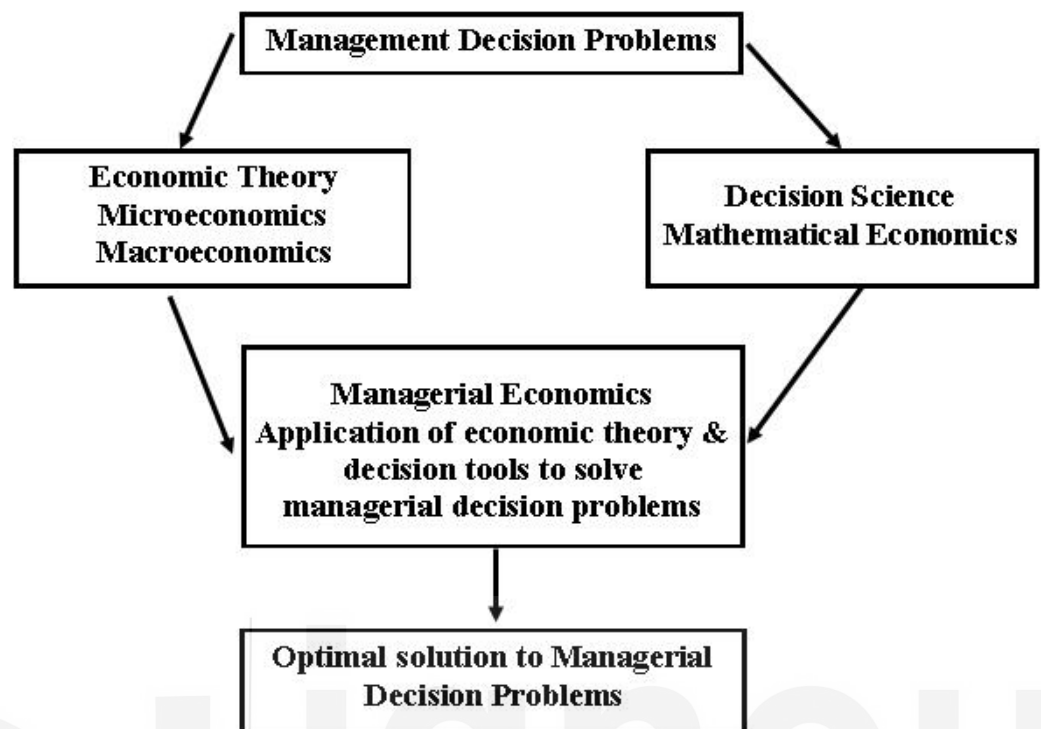
- It is concerned with “decision making of an economic nature.”
- It is “micro-economic” in character.
- It largely uses that body of economic concepts and principles, which is known as “theory of the firm.”
- It is “goal oriented and prescriptive”
- Managerial economics is both “conceptual and metrical”. It includes theory with measurement.

Figure 1.2: Decision-Making



Managerial economics should be thought of as applied microeconomics, which focuses on the behavior of the individual actors on the economic stage; firms and individuals.

Figure 1.3: Basic Characteristics



1.8 SUMMARY

Managerial economics is used by firms to improve their profitability. It is the economics applied to problems of choices and allocation of scarce resources by the firms. It refers to the application of economic theory and the tools of analysis of decision science to examine how an organisation can achieve its objective most efficiently. Managerial decisions are evaluated through concepts, tools and techniques of economic analysis of various types. It is linked with various fields of study.

1.9 SELF-ASSESSMENT QUESTIONS

1. Discuss the nature and scope of managerial economics.
2. “Managerial economics is the integration of economic theory with business practice for the purpose of facilitating decision-making and forward planning by manager”. Explain and comment.
3. Define scarcity and opportunity cost. What role do these two concepts play in the making of management decisions?
4. Managerial economics is often said to help the business student integrate the knowledge gained in other courses. How is this integration accomplished?
5. Compare and contrast microeconomics with macroeconomics. Although managerial economics is based primarily on microeconomics, explain why it is also important for managers to understand macroeconomics.
6. Justify that managerial economics is economics applied in decision-making.
7. What is the role of managerial economics in preparing managers?
8. How is managerial economics related to different disciplines?

1.10 FURTHER READINGS

Haynes, W.W. (1979). *Managerial Economics: Analysis and Cases* (3rd Ed.). Business Publications, Inc., Texas.

Adhikary, M. (1987). *Managerial Economics* (3rd Ed.). Khosla Publishers, Delhi.

Baumol, W.J. (1979). *Economic Theory and Operations Analysis* (4th Ed.). Prentice Hall India Pvt. Ltd., New Delhi.



UNIT 2 THE FIRM: STAKEHOLDERS, OBJECTIVES AND DECISION ISSUES

Objectives

After studying this unit, you should be able to:

- **understand** the rationale for existence of firms;
- **understand** the concept of economic profit and accounting profit;
- **appreciate** the use of opportunity cost;
- **differentiate** between various objectives of the firm.

Structure

- 2.1 Introduction
- 2.2 Objective of the Firm
- 2.3 Value Maximization
- 2.4 Alternative Objectives of the Firms
- 2.5 Goals of Real World Firms
- 2.6 Firm's Constraints
- 2.7 Basic Factors of Decision-Making: The Incremental Concept
- 2.8 The Equi-Marginal Principle
- 2.9 The Discounting Principle
- 2.10 The Opportunity Cost Principle
- 2.11 The Invisible Hand
- 2.12 Summary
- 2.13 Self-Assessment Questions
- 2.14 Further Readings

2.1 INTRODUCTION

The firm is an organisation that produces a good or service for sale and it plays a central role in theory and practice of Managerial Economics. In contrast to non-profit institutions like the 'Ford Foundation', most firms attempt to make a profit. There are thousands of firms in India producing large amount of goods and services; the rest are produced by the government and non-profit institutions. It is obvious that a lot of activities of the Indian economy revolve around firms.

One of the crucial determinants of a firm's behaviour is the state of technology. Technology imposes a limit on how much a firm can produce. It is the sum total of society's pool of knowledge concerning the industrial and agricultural arts. Production is any activity that transforms inputs into output and is applicable not only to the production of goods like steel and automobiles, but also to production of services like banking and insurance.

The firm changes hired inputs into saleable output. An input is defined as anything that the firm uses in its production process. Most firms require a wide array of inputs. For example, some of the inputs used by major steel firms like XYZ or ABC are iron ore, coal, oxygen, skilled labour of various

types, the services of blast furnaces, electric furnaces, and rolling mills as well as the services of the people managing the companies. To give another example, the inputs in production and sale of “chaat” by a street vendor are all the ingredients that go into making of the “chaat”, i.e. the stove, the “carrier”, and the services of the vendor. The inputs or the factors of production are divisible into two broad categories – human resources and capital resources. Labour resource and entrepreneurial resource are the two human resource inputs while land, man-made capital forests, rivers, etc. are the two capital resources. Thus the four major factors of production (FOP) are land, man-made capital, labour, and entrepreneur (organisation) while the remuneration they get is rent, interest (capital rental), wage, and profit, respectively.

The function of the firm, thus, is to purchase resources or inputs of labour services, capital and raw materials in order to convert them into goods and services for sale. There is a circular flow of economic activity between individuals and firms as they are highly interdependent. Labour has no value in the market unless there is a firm willing to pay for it. In the same way, firms cannot rationalize production unless some consumer is willing to buy their products. However, there is some incentive for each. Firms earn profits in turn satisfying the consumption demand of individuals and resource owners get wage, rent and interest payment. In the process of supplying the goods and services that consumers demand, firms provide employment to workers and also pay taxes that government uses to provide service (education, defence) that firms could not provide at all or as efficiently.

Essentially a firm exists because the total cost of production of output is lower than if the firm did not exist. There are several reasons for lower costs. Firstly, long term contract with labour saves the transaction costs because no new contract has to be negotiated every time a labour is to be hired or given new assignment. Secondly, there are government regulations like price-control and sales taxes also saved by having the transaction within the firm. Recall that sales tax is levied for transaction between firms and not within firms. When transactions take place within a firm they may be cheaper and hence such savings decrease the total cost of production of an output. In other words, the existence of firms could be explained by the fact that it saves transaction costs.

However, the size of the firm has to be limited because as the firms grow larger, a point is reached where the cost of internal transaction becomes equal to or greater than the cost of transaction between firms. When such a stage is reached, it puts a limit to the size of the firm. Further, the cost of supplying additional services like legal, medical etc. within the firm exceeds the cost of purchasing these services from other firms; as such services may be required occasionally.

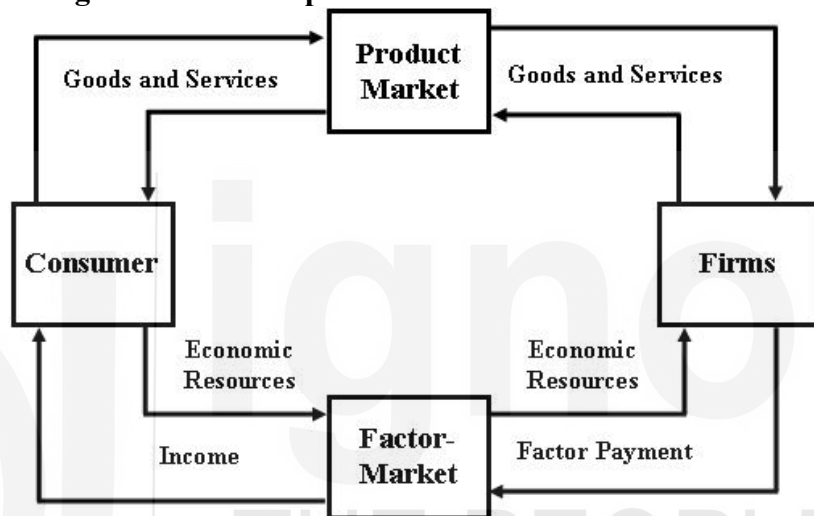
Let us consider the size of different kinds of firms around us and try to understand the reasons for such differences. Why are service firms generally smaller than capital-intensive firms like steel and automobile producing companies? What is the reason that a number of firms are choosing the BPO route? A part of the explanation must lie in the fact that it is cheaper to outsource than to absorb that activity within the firm. Consider a firm that needs to occasionally use legal service. Under what conditions will it choose

to hire a full time lawyer and take her on its rolls and under what conditions will the firm outsource the legal activity or hire legal services on a case-by-case basis. Naturally, the answer depends upon the frequency of use for legal services. The transaction cost framework demonstrates that the firm will contract out if the cost of such an arrangement is lower and will prefer in-house legal staff when the opposite is true.

Firms are classified into different categories as follows:

- a) Private sector firms.
- b) Public sector firms.
- c) Joint sector firms.
- d) Non-profit firms.

Figure 2.1: Interdependence of Consumers and Firms



Firms can also be classified on the basis of number of owners as:

- a) One Person Company.
- b) Partnership.
- c) Corporations

Some firms mentioned below are different from above. They may provide service to a group of clients for example, patients or to a group of its members only.

- a) Universities.
- b) Public Libraries.
- c) Hospitals.
- d) Museums.
- e) Churches.
- f) Voluntary Organisations.
- g) Cooperatives.
- h) Unions.
- i) Professional Societies, etc.

The concept of a firm plays a central role in the theory and practice of managerial economics. It is, therefore, valuable to discuss the objectives of a firm.

2.2 OBJECTIVE OF THE FIRM

The traditional objective of the firm has been profit maximization. It is still regarded as the most common and theoretically the most plausible objective of business firms. We define profits as revenues less costs. But the definition of cost is quite different for the economist than for an accountant. Consider an independent businessperson who has an MBA degree and is considering investing ₹1 lakh in a retail store that s/he would manage. There are no other employees. The projected income statement for the year as prepared by an accountant is as shown below:

Sales			₹90,000
Less: Cost of Goods Sold	₹40,000		
Gross Profit:			₹50,000
Less:			
	Advertising	₹10,000	
	Depreciation	₹10,000	
	Utilities	₹3,000	
	Property Tax	₹2,000	
	Misc.	₹5,000	
			=₹30,000
Net Accounting Profit			₹20,000

This accounting or business profit is what is reported in publications and in the quarterly and annual financial reports of businesses.

The economist recognizes other costs, defined as implicit costs. These costs are not reflected in cash outlays by the firm, but are the costs associated with foregone opportunities. Such implicit costs are not included in the accounting statements but must be included in any rational decision making framework. There are two major implicit costs in this example. First, the owner has ₹1 lakh invested in the business. Suppose the best alternative use for the money is a bank account paying a 10 per cent interest rate. This risk less investment would return ₹10,000 annually. Thus, ₹10,000 should be considered as the implicit or opportunity cost of having ₹1 lakh invested in the retail store.

Let us consider the second implicit cost, which includes the manager's time and talent. The annual wage return on an MBA degree may be taken as ₹35,000 per year. This is the implicit cost of managing this business rather than working for someone else. Thus, the income statement should be amended in the following way in order to determine the economic profit:

Sales			₹90,000
Less: Cost of Goods Sold	₹40,000		
Gross Profit:			₹50,000
Less:			
	Advertising	₹10,000	
	Depreciation	₹10,000	
	Utilities	₹3,000	
	Property Tax	₹2,000	
	Misc.	₹5,000	
			=₹30,000
Accounting Profit		₹20,000	

Less: Implicit Costs:			
	Return on ₹ 1 lakh of capital	₹10,000	} = ₹45,000
	Foreign Wages	₹35,000	
Net "Economic Profit"		₹25,000	

Looking at this broader perspective, the business is projected to lose ₹25,000 in the first year. ₹ 20,000 accounting profit disappears when all relevant costs are included. Another way of looking at the problem is to assume that ₹1 lakh had to be borrowed at, say, 10 per cent interest and an MBA graduate hired at ₹35,000 per year to run the store. In this case, the implicit costs become explicit and the accounting made explicit. Obviously, with the financial information reported in this way, an entirely different decision might be made on whether to start this business or not.

Thus, we can say that economic profit equals the revenue of the firm minus its explicit costs and implicit costs. To arrive at the cost incurred by a firm, a value must be put to all the inputs used by the firm. Money outlays are only a part of the costs. As stated above, economists also define opportunity cost. Since the resources are limited, and have alternative uses, you must sacrifice the production of a good or service in order to commit the resource to its present use. For example, if by being the owner manager of your firm, you sacrifice a job that offers you ₹ 2,00,000 per annum, then two lakhs is your opportunity cost of managing the firm. Similarly, if Suresh Kumar was not playing cricket, he could have earned a living (perhaps, not such a good one!) by being a cricket commentator. Suresh's opportunity cost of playing cricket is the amount he could have earned being a cricket commentator.

The assignment of monetary values to physical inputs is easy in some cases and difficult in others. All economic costing is governed by the principle of opportunity cost. If the firm maximizes profits, it must evaluate its costs according to the opportunity cost principle. Assigning costs is straightforward when the firm buys an input on a competitive market. Suppose the firm spends ₹ 20,000 on buying electricity. For its factory, it has sacrificed claims to whatever else Rs 20,000 can buy and thus the purchase price is a reasonable measure of the opportunity cost of using that electricity. The situation is the same for hired factors of production. However, a cost must be assigned to factors of production that the firm neither purchases, nor hires because it already owns them. The cost of using these inputs is implicit costs and has to be imputed. Implicit costs arise because the alternative (opportunity) cost doctrine must be applied to the firm. The profit calculated after including implicit as well as explicit costs in total cost is called economic profit.

Profit plays two primary roles in the free-market system. First, it acts as a signal to producers to increase or decrease the rate of output, or to enter or leave an industry. Second, profit is a reward for entrepreneurial activity, including risk taking and innovation. In a competitive industry, economic profits tend to be transitory. The achievement of high profits by a firm usually results in other firms increasing their output of that product, thus reducing price and profit. Firms that have monopoly power may be able to earn above-normal profits over a longer period; such profit does not play a socially useful role in the economy.

Although, profit maximization is a dominant objective of the firm, other important objectives of the firm, other than profit maximization that we will discuss in this unit are:

1. Maximization of sales revenue.
2. Maximization of firm's growth rate
3. Maximization of manager's own utility or satisfaction
4. Making a satisfactory rate of profit.
5. Long-run survival of the firm
6. Entry-prevention and risk avoidance.

Activity 3

- a) "Among the various objectives of a modern firm, profit maximization is the most important". Comment.

.....

.....

.....

.....

- b) Outline the circular flow of economic activity between individuals and firms.

.....

.....

.....

.....

- c) (i) profit is a cost of doing business and is the amount by which exceeds profit.

(ii) When a firm earns just a normal rate of return, equals total economic cost and profit is zero.

- d) A firm collects ₹ 1.75 lakhs in revenue and spends ₹ 80 thousands on raw materials in a year. The owners of the firm have provided ₹ 5 lakhs of their own money to the firm instead of investing the money and earning a high rate of interest.

- (i) The firm earns economic profit of The firm's normal profit is
- (ii) The firm's accounting profit is
- (iii) If the firm's costs stay the same but its revenue falls to only a normal profit is earned.

2.3 VALUE MAXIMIZATION

Most firms have sidelined short-term profit as their objective. Firms are often found to sacrifice their short-term profit for increasing the future long-term profit. Thus, the theory states that the objective of a firm is to maximize wealth or value of the firm. For example, firms undertake research and development expenditure, expenditure on new capital equipment or major marketing programmes which require expenditure initially but are meant to

generate future profits. The objective of the firm is thus to maximize the present or discounted value of all future profits and can be stated as:

$$PV(\pi) = \frac{\pi_1}{(1+r)^1} + \frac{\pi_2}{(1+r)^2} + \dots + \frac{\pi_n}{(1+r)^n}$$

$$= \sum_{t=1}^n \frac{\pi_t}{(1+r)^t}$$

Where, PV = Present Value of all expected future profits of the firm.

$\pi \dots \pi_n$ = Expected profit in 1, 2,n years.

r = Appropriate discount rate

t = Time period 1n.

Assumed profit is equal to total revenue (TR) minus total cost (TC), then the value

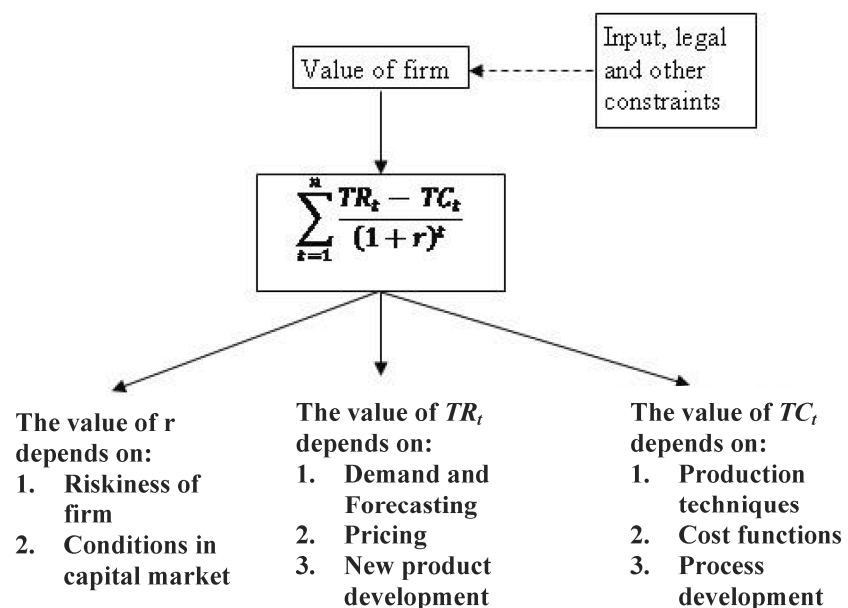
of the firm can also be stated as:

$$\text{Value of the firm} = \sum_{t=1}^n \frac{TR_t - TC_t}{(1+r)^t}$$

Thus maximizing the discounted value of all future profits is equivalent to maximizing the value of the firm.

A careful inspection of the equation suggests how a firm's managers and workers can influence its value. For example, in a company, the marketing managers and sales representatives work hard to increase its total revenues, while its production managers and manufacturing engineers strive to reduce its total costs. At the same time, its financial managers play a major role in obtaining capital, and hence influence the equation, while its research and development personnel invent and reduce its total costs. All of these diverse groups affect the company's value, defined here as the present value of all expected future profits of the firm.

Figure 2.2: Determination of the value of the firm



2.4 ALTERNATIVE OBJECTIVES OF THE FIRMS

Economists have also examined other objectives of firms. We shall discuss some of them here. According to Baumol, most managers will try to maximize sales revenue. There are many reasons for this. For example, the salary and other earnings of managers are more closely related to sales revenue than to profits. Banks and financiers look at sales revenue while financing the corporation. The sales revenue trend is a readily available indicator of performance of the firm. Growth in sales increases the competitive strength of the firm. However, in the long run, sales maximization and profit maximization may converge into one objective.

Another economist Robin Marris assumes that owners and managers have different utility functions to maximize. The manager's utility function (U_m) and Owner's utility functions (U_o) are:

$$U_m = f(\text{Salary, job, power, prestige, status})$$

$$U_o = f(\text{Output, capital, profit, share})$$

By maximizing the variables, managers maximize both their own utility function and that of the owners. Most of the variables of both managers and owners are correlated with a single variable, namely, the size of the firm. Maximization of these variables depends on the growth rate of the firm. Thus, Marris argues that managers will attempt to maximize growth rate of firms. However, this objective does not completely discard the profit maximization objective.

According to Oliver Williamson, managers seek to maximize their own utility function subject to a minimum level of profit. The utility function which managers seek to maximize include both quantifiable variables like salary and slack earnings and non-quantifiable variables like power, status, security of job, etc. The model developed by Cyert-March focuses on satisficing behaviour of managers. The firm has to deal with an uncertain business world and managers have to satisfy a variety of groups-staff, shareholders, customers, suppliers, authorities, etc. All these groups have often-conflicting interests in the firm. In order to reconcile between the conflicting interests and goals, managers form an aspiration level of the firm combining the following objectives – production, sales and market share, inventory and profit. The aspiration levels are modified and revised on the basis of achievements and changing business environment.

As is true with most economic models, the application will depend upon the situation and one cannot say that a particular model is better than the other. In general, one can assert that the profit maximising assumption seems to be a reasonable approximation of the real world, although in certain cases there might be a deviation from this objective.

2.5 GOALS OF REAL WORLD FIRMS

By now we know that firms that maximize profits are not just concerned about short-run profits, but are more concerned with long-term profits. They may not take full advantage of a potential monopolistic situation, for example, many stores have liberal return policies; many firms spend millions on improving their reputation and want to be known as 'good' citizens. The decision maker's income is often a cost of the firm. Most real-world production takes place in large corporations with 8-9 levels of management, thousands of stakeholders and boards of directors. Self-interested decision makers have little incentive to hold down their pay. If their pay is not held down, firm's profit will be lower. Most firms manage to put some pressure on managers to make at least a pre-designated level of profit.

In the modern corporation, the owners or stakeholders (i.e. the principals) hire managers (i.e., agents) to conduct the day-to-day operations of the firm. These managers are paid a salary to represent the interest of the owners, ostensibly, to maximize the value of the firm. A board of directors is elected by the owners to meet regularly with the managers to oversee their activity and to try to ensure that the managers are, in fact, acting in the best interest of the owners.

Because of the difficulty of monitoring the managers on a continual basis, it is possible that goals other than profit-maximization may be pursued. In addition to those mentioned earlier, the managers may seek to enhance their positions by spending corporate funds on fancy offices, excessive and expensive travel, club memberships, and so forth. In recent years, many corporations have taken action to align the interests of owners with the interests of the managers by tying a large share of managerial compensation to the financial performance of the firm.

For example, the manager may be given a basic salary plus potentially large bonuses for meeting such goals as attaining a specified return on capital, growth in earnings, and/or increase in the price of the firm's stock. With regard to the latter, the use of stock options awarded to top managers is a most effective way to ensure that managers act in the interest of the shareholders. Typically, the arrangement provides that the manager is to receive an option to buy a specified number of shares of common stock at the current market price for a specified number of years. The only way the executives can benefit from such an arrangement is if the price of stock rises during the specified term. The option is exercised by buying the shares at the specified price, and the gain equals the increase in share price multiplied by the number of shares purchased. Sometimes the agreement specifies that the stock must be held for several years following purchase. Essentially, this option arrangement makes the manager a de facto owner, even if the option has not been exercised. In almost every case of a report of unusually high executive compensation, the largest part of that compensation is associated with gains from stock options.

Emergence of oligopoly, a market structure characterized by the existence of a few large firms, mergers and amalgamations have made the structure of industries concentrated so that few large (dominant) firms account for a major portion of an industry's output. This shifts the pressure on each firm to

maximize profit independently and leads to joint profit maximizations through cartels and collusions. Profit maximization may not be the only inevitable objective.

India's Global Companies and their Objectives: One of the most significant business and economic trends of the late twentieth century is the rise of 'global' or 'stateless' corporation. The trends toward global companies are unmistakable and are accelerating. The sharpest weapon that a corporation can develop to survive and thrive, in the globalised market place is competitiveness. Its corner stone as articulated by strategy guru Michael Porter is its ability to create more value on a sustainable basis, for the customer than its rivals can.

Many Indian corporations such as an Indian multinational conglomerate Pharmaceutical company among others are competing on the world stage. Whatever product or service a company offers it must meet the customers wants in the most satisfactory manner. This should be the aim of the company. The competitiveness of Company in the global market place comes from both quality and scale. The challenge is to remain at the top. That challenge is linked with productivity. Pharmaceutical Company's greatest strength lies in the fact that it is strongly backward integrated. It helps them manage cost across the entire value chain making them extremely cost competitive. Cost leadership is a function of scale and technology. By upgrading technology, Pharmaceutical Company's could continue to be a cost leader. A company has to continuously upgrade itself on several parameters: production efficiency, product development, quality management and marketing skills.

This competitiveness - defined by Michael Porter as the sustained ability to generate more value for customers than the cost of creating that value. Many new entrepreneurs and new industries are emerging who are able to operate successfully in this changed environment

2.6 FIRM'S CONSTRAINTS

Decision-making by firms takes place under several restrictions or constraints, such as:

Resource Constraints: Many inputs may be available in a limited or fixed quantity e.g., skilled workers, imported raw material, etc.

Legal Constraints: Both individuals and firms have to obey the laws of the State as well as local laws. Environmental laws, employment laws, disposal of wastes are some examples.

Moral Constraints: These imply to actions that are not illegal but are sufficiently consistent with generally accepted standards of behaviour.

Contractual Constraints: These bind the firm because of some prior agreement such as a long-term lease on a building or a contract with a labour union that represents the firm's employees.

Decision-making under these constraints with optimal results is a fundamental part of managerial economics.

2.7 BASIC FACTORS OF DECISION-MAKING: THE INCREMENTAL CONCEPT

Incremental reasoning involves estimating the impact of decision alternatives. The two basic concepts in the incremental analysis are:

Incremental Cost (IC)
Incremental Revenue (IR)

Incremental cost is defined as the change in total cost as a result of change in the level of output, investment etc. Incremental revenue is defined as the change in total revenue resulting from a change in the level of output, prices etc. A manager always determines the worth of a decision on the basis of the criterion that $IR > IC$.

A decision is profitable if

- it increases revenue more than it increases cost
- it reduces some costs more than it increases others
- it increases some resources more than it decreases others
- it decreases costs more than it decreases revenues.

To illustrate the above points, let us take a case where a firm gets an order that can get it additional revenue of ₹ 2,000. The normal cost of production of this order is—

Labour	: ₹ 600
Materials	: ₹ 800
Overheads	: ₹ 720
Selling and administration expenses	: ₹ 280
Full cost	: ₹ 2,400

Comparing the additional revenue with the above cost suggests that the order is unprofitable. But, if some existing facilities and underutilized capacity of the firm were utilized, it would add much less to cost than ₹ 2,400. For example, let us assume that the addition to cost due to this new order is, say, the following:

Labour	: ₹ 400
Materials	: ₹ 800
Overheads	: ₹ 200
Total incremental Cost	: ₹ 1,400

In the above case the firm would earn a net profit of ₹ 2000 – ₹ 1400 = ₹ 600, while at first it appeared that the firm would make a loss of ₹ 400 by accepting the order.

The worth of such a decision can be judged on the basis of the following theorem.

Theorem I: A course of action should be pursued upto the point where its incremental benefits equal its increment costs.

According to the theorem, the firm represented in Table 2.1 will produce only seven units of output as its Marginal Revenue (MR)= Marginal Cost (MC)¹ at that level of output. As can be calculated from the Table, the MC of 8th unit is more than its MR. Hence the firm gets negative profit from 8th unit and thus is advised not to produce it.

The acceptance or rejection of an order by a firm for its product depends on whether the resultant costs are greater or less than the resultant revenue. If these principles are not followed, the equilibrium position would be disturbed. But the problem with the concept of marginalism is that the independent variable may be subject to “bulk changes” instead of “unit changes”. For example, a builder may not change one labourer at a time, but many of them together. Similarly, the output may change because of a change in process, pattern or a combination of factors, which may not always be measured in unit terms. In such cases, the concept of marginalism is changed to incrementalism. Or, in other words, incrementalism is more general, whereas marginalism is more specific. All marginal concepts are incremental concepts, but all incremental concepts need not be marginal concepts.

Table 2.1 : Profit Function of a Firm

Unit of Output	Total Revenue	Total Cost	Total Profit	Average Profit	Marginal Profit
1	2	3	4	5	6
1	20	15	5	5.0	-
2	40	29	11	5.5	6
3	60	42	18	6.0	7
4	80	52	28	7.0	10
5	100	65	35	7.0	7
6	120	81	39	6.5	4
7	140	101	39	5.6	0
8	160	125	35	4.4	-4

2.8 THE EQUI-MARGINAL PRINCIPLE

According to this principle, different courses of action should be pursued upto the point where all the courses provide equal marginal benefit per unit of cost. It states that a rational decision-maker would allocate or hire his resources in such a way that the ratio of marginal returns and marginal costs of various uses of a given resource or of various resources in a given use is the same. For example, a consumer seeking maximum utility (satisfaction)

¹ Marginal Revenue is the additional revenue from selling one more unit, while Marginal Cost is the additional cost of producing one more unit.

from his consumption basket, will allocate his consumption budget on goods and services such that

$$\begin{aligned} MU_1/MC_1 &= MU_2/MC_2 = \dots\dots\dots = MU_n/MC_n \\ \text{Where } MU_1 &= \text{marginal utility from good one,} \\ MC_1 &= \text{marginal cost of good one and so on,} \end{aligned}$$

Similarly, a producer seeking maximum profit would use the technique of production (input-mix.) which would ensure

$$MRP_1/MC_1 = MRP_2/MC_2 = \dots\dots\dots MRP_n/MC_n$$

Where MRP_1 = Marginal revenue product of input one (e.g. Labour),
 MC_1 = Marginal cost of input one and so on.

It is easy to see that if the above equation was not satisfied, the decision makers could add to his utility/profit by reshuffling his resources/input e.g. if $MU_1/MC_1 > MU_2/MC_2$ the consumer would add to his utility by buying more of good one and less of good two. Table 2.2 summarizes this principle for different sellers.

Example: A multi-commodity consumer wishes to purchase successive units of A, B and C. Each unit costs the same and the consumer is determined to have a combination including all the three items. His budget constraint is such that he cannot buy more than six units in all. Again, he is subject to diminishing marginal utility i.e. as he has more of an item, he wants to consume less of it. Table 2.3 shows the optimization example:

Table 2.2: The Equi-Marginal Principle

Unit	Equi-Marginal Principle
Multi-market seller	$MR_1 = MR_2 = MR_3 = \dots\dots\dots MR_n$
Multi-plant monopolist	$MC_1 = MC_2 = MC_3 = \dots\dots\dots MC_n$
Multi-factor employer	$MP_1 = MP_2 = MP_3 = \dots\dots\dots MP_n$
Multi-product firm	$M\pi_1 = M\pi_2 = M\pi_3 = \dots\dots\dots M\pi_n$
Multi-commodity consumer	$MU_1 = MU_2 = MU_3 = \dots\dots\dots MU_n$

MR=marginal revenue; MC=marginal cost; MP=marginal product;
M π =marginal profit; MU=marginal utilities.

Table 2.3 : Optimization Example

Units	MARGINAL UTILITES		
	Item A	Item B	Item C
1	10	9	8
2	9	8	7
3	8	7	6
4	7	6	5
5	6	5	4
6	5	4	3

The utility maximizing consumer will end up with a purchase of $3A+2B+1C$ because that combination satisfies equi-marginalism:

$$MU_a = MU_b = MU_c = 8$$

In the real world, often the equi-marginalism concept has to be replaced by equi-incrementalism. This is because, changes in the real world are discrete or lumpy and therefore the concept of marginal change may not always apply. Instead, changes will be incremental in nature, but the decision rule or optimizing principle will remain the same

2.9 THE DISCOUNTING PRINCIPLE

Many transactions involve making or receiving cash payments at various future dates. A person who takes a house loan trades a promise to make monthly payments for say, fifteen or twenty years for a large amount of cash now to pay for a home. This case and other similar cases relate to the time value of money. The time value of money refers to the fact that a rupee to be received in the future is not worth a rupee today. Therefore, it is necessary to have techniques for measuring the value today (i.e., the present value) of rupees to be received or paid at different points in the future. This section outlines the approach to analyzing problems that involve payment and/or receipt of money at one or more points in time.

One may ask how much money today would be equivalent to ₹ 100 a year from now if the rate of interest is 5%. This involves determining the present value of ₹ 100 to be received after one year. Applying the formula –

$$PV = \frac{100}{1.05}$$

we obtain ₹ 95.24,

₹ 95.24 will accumulate to an amount exactly equal to ₹ 100 in one year at the interest rate of 5 per cent. Looked at another way, you will be willing to pay maximum of ₹ 95.24 for the benefit of receiving ₹ 100 one year from now if the prevailing interest rate is 5 per cent.

The same analysis can be extended to any number of periods. A sum of ₹ 100 two years from now is worth:

$$PV_2 = \frac{100}{1.05^2}$$

= ₹90.70 today.

In general, the present value of a sum to be received at any future date can be found by the following formula:

$$PV = \frac{R_n}{(1 + i)^n}$$

PV = present value, R_n = amount to be received in future, i = rate of interest, n = number of years lapsing between the receipt of R .

If the receipts are made available over a number of years, the formula becomes:

$$PV = \frac{R_1}{1+i} + \frac{R_2}{(1+i)^2} + \frac{R_3}{(1+i)^3} + \dots + \frac{R_n}{(1+i)^n}$$

$$PV = \sum_{k=1}^n \frac{R_k}{(1+i)^k}$$

In the above formula if $R_1 = R_2 = R_3$ etc., it becomes an 'annuity'. An annuity has been defined as series of periodic equal payments. Although the term is often thought of in terms of a retirement pension, there are many other examples of annuities. The repayment schedule for a home loan is an annuity. A father's agreement to send his son ₹ 1000 each month while he is in college is another example. Usually, the number of periods is specified, but not always. Sometimes retirement benefits are paid monthly as long as a person is alive. In other case, the annuity is paid forever and is called 'perpetuity.'

It must be emphasized that the strict definition of an annuity implies equal payments. A contract to make 20 annual payments, which increase each year by, say, 10 per cent, would not be an annuity. As some financial arrangements provide for payments with periodic increase, care must be taken not to apply an annuity formula if the flow of payments is not a true annuity.

The present value of an annuity can be thought of as the sum of the present values of each of several amounts. Consider an annuity of three ₹ 100 payments at the end of each of the next three years at 10 percent interest. The present value of each payment is

$$PV_1 = 100 \frac{1}{1.10}$$

$$PV_2 = 100 \frac{1}{1.10^2}$$

$$PV_3 = 100 \frac{1}{1.10^3}$$

and the sum of these would be

$$PV = 100 \frac{1}{1.10} + 100 \frac{1}{(1.10)^2} + 100 \frac{1}{(1.10)^3}$$

OR

$$PV = 100 \left[\frac{1}{1.10} + \frac{1}{(1.10)^2} + \frac{1}{(1.10)^3} \right]$$

The present value of this annuity is

$$PV = 100 (0.9091 + 0.8264 + 0.7513) = 100 (2.4868) = 248.68$$

Although this approach works, it clearly would be cumbersome for annuities of more than a few periods. For example, consider using this method to find the present value of a monthly payment for forty years if the monthly interest rate is 1 per cent. That would require evaluating the present value of each of 480 amounts. In general, the formula for the present value of an annuity of A rupees per period for n periods and a discount rate of i is

$$PV = A \frac{1}{(1+i)} = A \frac{1}{(1+i)^2} + \dots + A \frac{1}{(1+i)^n}$$

2.10 THE OPPORTUNITY COST PRINCIPLE

The opportunity cost of anything is the return that can be had from the next best alternative use. A farmer who is producing wheat can also produce potatoes with the same factors. Therefore, the opportunity cost of a quintal of wheat is the amount of the output of potatoes given up. The opportunity costs are the ‘costs of sacrificed alternatives.’

Whenever the manager takes a decision he chooses one course of action, sacrificing the other alternative courses. We can therefore evaluate the one, which is chosen in terms of the other (next best) alternative that is sacrificed. A machine can produce either X or Y. The opportunity cost of producing a given quantity of X is the quantity of Y which it would have produced.

The opportunity cost of holding ₹1000 as cash in hand for one year is the 10% rate of interest, which would have been earned had it been invested in the form of fixed deposits in the bank.

- all decisions which involve choice must involve opportunity cost calculation,
- the opportunity cost may be either real or monetary, either implicit or explicit, either non-quantifiable or quantifiable.

Opportunity costs’ relevance is not limited to individual decisions. Opportunity costs are also relevant to government’s decisions, which affect everyone in society. A common example is the guns-versus-butter debate. The resources that a society has are limited; therefore its decisions to use those resources to have more guns (more weapons) means that it must have less butter (fewer consumer goods). Thus when society decides to spend 100 crore on developing a defence system, the opportunity cost of that decision is 100 crores not spent on fighting drugs, helping the homeless, or paying off some of the national debt.

For the country as a whole, the production possibility reflects opportunity costs. Figure 2.1 shows the Production Possibility Curve (PPC) reflecting the different combinations of goods, which an economy can produce, given its state of technology and total resources. It illustrates the menu of choices open to the economy. Let us take the example that the economy can produce only two goods, butter and guns. The economy can produce only guns, only butter or a combination of the two, illustrating the tradeoffs or choice inherent in such a decision. The opportunity cost of choosing guns over butter increases as the production of guns is increased. The reason is that some resources are relatively better suited to producing guns. The quantity of butter, which has to be sacrificed to produce an additional unit of guns, is called the

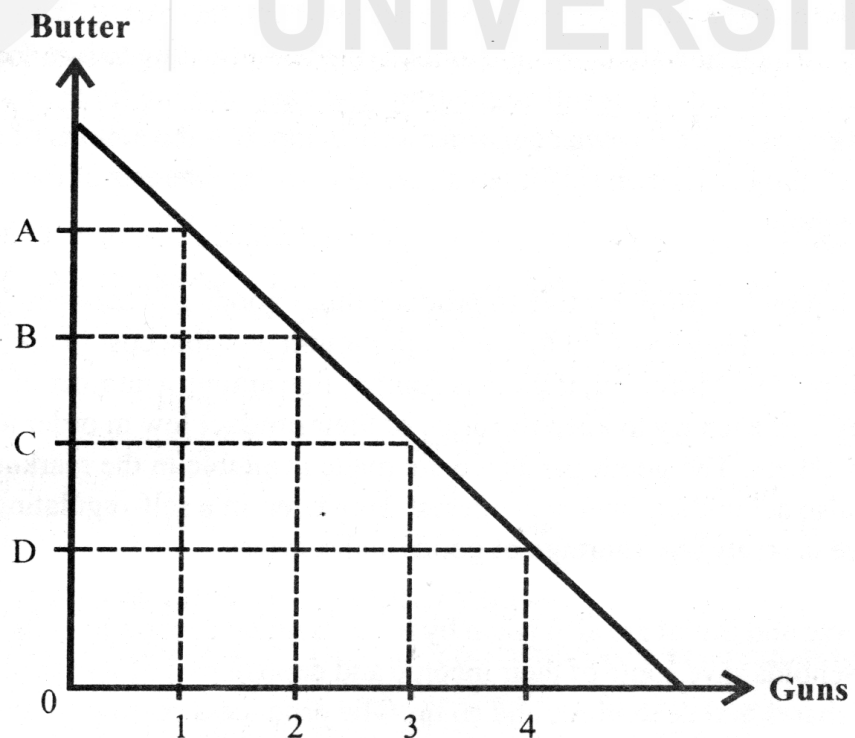
opportunity cost of guns (in terms of butter). Due to the increasing opportunity cost of guns, the PPC curve will be concave to the origin. Increasing opportunity cost of guns means that to produce each additional unit of guns, more and more units of butter have to be sacrificed. The basis for increasing opportunity costs is the following assumptions:

- i) Some factors of production are more efficient in the production of butter and some more efficient in production of guns. This property of factors is called specificity. Thus specificity of factors of production causes increasing opportunity costs.

Figure 2.1: Production Possibility Curve



Figure 2.2: Production Possibility Curve - Linear



- ii) The production of the goods require more of one factor than the other. For example, the production of guns may require more capital than that of butter. Hence, as more and more of capital is used in the manufacture of guns, the opportunity cost of guns is likely to increase.

Let us assume that an economy is at point A where it uses all its resources in the production of butter. Starting from A, the production of 1 unit of guns requires that AC units of butter be given up. The production of a second unit of guns requires that additional CD units of butter be given up. A third requires that DE be given up, and so on. Since $DE > CD > AC$, and so on, it means that for every additional unit of guns more and more units of butter will have to be sacrificed, or in other words, the opportunity cost keeps on increasing.

The opportunity cost of the first few units of guns would initially be low and those resources, which are more efficient in the production of guns move from, butter production to gun production. As more and more units of guns are produced, however, it becomes necessary to move into gun production, even for those factors, which are more efficient in the production of butter. As this happens, the opportunity cost of guns gets larger and larger. Thus, due to increasing opportunity costs the PPC is concave.

If the PPC curve were to be a straight line, the opportunity cost of guns would always be constant. This would mean equal (and not increasing amounts of butter) would have to be forgone to produce an additional unit of guns. The assumption of constant opportunity costs is very unrealistic. It implies that all the factors of production are equally efficient either in the production of butter or in the production of guns.

For many of the choice society make opportunity costs tend to increase as we choose more and more of an item. Such a phenomenon about choice is so common, in fact, that it has acquired a name: the principle of increasing marginal opportunity cost. This principle states that in order to get more of something, one must give up ever-increasing quantities of something else. In other words, initially the opportunity costs of an activity are low, but they increase the more we concentrate on that activity.

2.11 THE INVISIBLE HAND

Adam Smith, the father of modern economics believed that there existed an “invisible hand” which ruled over the economic system. According to him the economic system, left to itself, is self-regulating. The basic driving force in such a system is trying to enhance its own economic well-being. But the actions of each unit, acting according to its own self-interest, are also in the interests of the economy as a whole.

Producers are led by the profit motive to produce those goods and services which the consumers want. They try to do this at the minimum possible cost in order to maximize their profits. Moreover, if there is competition among a number of producers, they will each try to keep the price of their product low in order to attract the consumers. The goods produced are made available in the market by traders. They also act in their own self-interest. However, in a self-regulating economy, there is rarely any shortage of goods and services.

Decisions to save and invest are also taken by the individual economic units. For example, households save some of their income and deposit part of it in the banks, or invest it in shares and debentures and so on. The producers borrow from the banking system and also issue shares and debentures to finance their investments. In turn, they reinvest a part of their profits.

All the economic functions have been carried out by individuals acting in isolation. There is no government or centralized authority to determine who should produce what and in what quantity, and where it should be made available. Yet in a self-regulating economy there is seldom a shortage of goods and services. Practically everything you want to buy is available in the market. Thus according to Adam Smith, the economic system is guided by the “invisible hand”. In a more technical way we can say that the basic economic problems in a society are solved by the operation of market forces.

2.12 SUMMARY

There is a circular flow of economic activity between individuals and firms as they are highly interdependent. The firms’ existence is based on manifold reasons. Firms are classified into different categories. Different firms belonging to the same industry, facing the same market environment, behave differently. Thus, the necessity for theories of the firm. Profit is defined as revenues minus costs. But the definition of cost is quite different for economist than for the accountant. Short-term profit has been sidelined by most firms as their objective for increasing the future long-term profit. Real world firms often have a set of complicated goals. The basic factors of decision making can be outlined by various principles.

2.13 SELF-ASSESSMENT QUESTIONS

1. Write notes in about 200 words on the following:
 - a) The incremental concept
 - b) Opportunity cost
 - c) Scope of managerial economics
 - d) The Invisible Hand
2. ‘Managerial Economics serves as a link between traditional economics and decision sciences for business decision-making.’ Elucidate.
3. Calculate, using the best estimates you can make:
 - a) Your opportunity cost of attending college.
 - b) Your opportunity cost of taking this course.
 - c) Your opportunity cost of attending yesterday’s lecture of your course.
4. The following is the hypothetical production possibility table of India:

Resources Devoted to Clothing	Output of Clothing	Output of Food
100%	20	0
80%	16	5
60%	12	9
40%	8	12
20%	4	14
0%	0	15

The Firm: Stakeholders, Objectives and Decision Issues

- Draw India's production possibility curve.
- What is happening to marginal opportunity costs as output of food increases?
- If the country gets better at the production of food, what will happen to the production possibility curve?

If the country gets equally better at producing food and producing clothing, what will happen to the production possibility curve?

- Use the following interest rates for government bonds for the risk-free discount rate and answer the following:

Time of Maturity (years)	Interest Rate (%)
1	5.75
2	6.00
3	6.25
4	6.50
5	6.75

- Calculate the PV of a ₹ 1 lakh payment to be received at the end of one year, 2 years, 3 years, 4 years and 5 years.
- What is the present value of a firm with a 5 years life span that earns the following stream of expected profit at the year-end?

Years	Expected Profit (in Crores)
1	10
2	20
3	50
4	25
5	50

- Value maximization has become the major objective of a modern firm. Comment.

2.14 FURTHER READINGS

Dean, J. (1951). *Managerial Economics*. PrenticeHall.

Mote, V.L., Paul, S., & Gupta, G.S. (2016). *Managerial Economics Concepts and Cases*, Tata McGraw-Hill, New Delhi.

Koutsoyiannis, A. (2018). *Modern Microeconomics* (2nd Ed.). Macmillan Publishers India Pvt. Ltd.

UNIT 3 BASIC CONCEPTS & TECHNIQUES

Objectives

After studying this unit, you should be able to:

- **identify** a wide range of techniques used in managerial economics;
- **apply** the techniques to understand the meaning of the results;
- **identify** the strengths and weaknesses of the different meth.

Structure

- 3.1 Introduction
- 3.2 Opportunity Set
- 3.3 Variables and Constants
- 3.4 Derivatives
- 3.5 Partial Derivatives
- 3.6 Optimization Concept
- 3.7 Regression Analysis
- 3.8 Specifying the Regression Equation
- 3.9 Estimating the Regression Equation
- 3.10 Decision Under Risk
- 3.11 Uncertainty Analysis and Decision Making
- 3.12 Role of Managerial Economist
- 3.13 Summary
- 3.14 Self-Assessment Questions
- 3.15 Further Readings

3.1 INTRODUCTION

The manager needs various techniques to assist and help him in making decisions that will ultimately maximize the value of the firm. These techniques and tools are quantitative in nature. The introduction of some commonly used tools used in managerial decision making becomes imperative.

In this unit we are going to discuss some basic techniques which would be helpful in understanding the concept of managerial economics, in turn helping us to apply these techniques as and when required.

3.2 OPPORTUNITY SET

A set is a collection of distinct or well defined objects like (5, 6, 7) or (a, b, c). For example listing of all residents of Delhi or all animals in a zoo is difficult. Thus a set is also formed by developing a criterion for membership. For example the set of all positive numbers between 1 and 10 or set of all points lying on the line $x + y = 4$.

In managerial economics the need is to define an opportunity set of a decision maker, i.e., the set of alternative actions which are feasible. For example, the opportunity set of a consumer is the set of all combinations of goods which

the consumer can buy with his given income. Given the consumer's budget and prices of all goods, the opportunity set is well defined, and we can find out whether the consumers can buy that combination of goods. If a person consumes only two goods (x , y), whose prices are Rs. 5 and Rs. 4 per unit, and his income is Rs. 100, then the opportunity set is as is shown in Figure 3.1.

Figure 3.1: Opportunity Set



If the consumer spends all his income on X, he can buy 20 units of X with no Y (B). or if he spends all on good Y, then he buys 25 units of Y and no X (A). All other alternatives of spending Rs. 100 will lie on the line AB. Hence the area OAB constitutes the consumer's opportunity set.

3.3 VARIABLES AND CONSTANTS

Variables are a measure which change and can take a set of possible values within a given problem. A constant or parameter is a quantity which does not change in a given problem. For example, consider $Y=a+bX$; here, a and b are constants and X and Y are variables. X is the **independent variable** or **exogenous variable** while Y is the **dependent** or **endogenous variable**. The value of X will be given from outside the system, while values of Y will be determined from within the system (by the equation). X can assume different values and this will cause Y to assume different values also.

We can show the relationship between two variables Q and P by a function, such that for every value of P there is only one value of Q . In this case of function is $Q = Q(P)$. These are the basic building blocks of economic models. The function $Q=Q(P)$ is a demand function and its graph which price on one and quantity on the other axis will give a demand curve. P , is the argument of the function or the independent variable and Q is the dependent' variable. It indicates the cause-effect relationship between variables. (P is the cause variable, while Q is the effect variable).

A function can be represented by means of a table or graph.

Figure 3.2: Linear Function

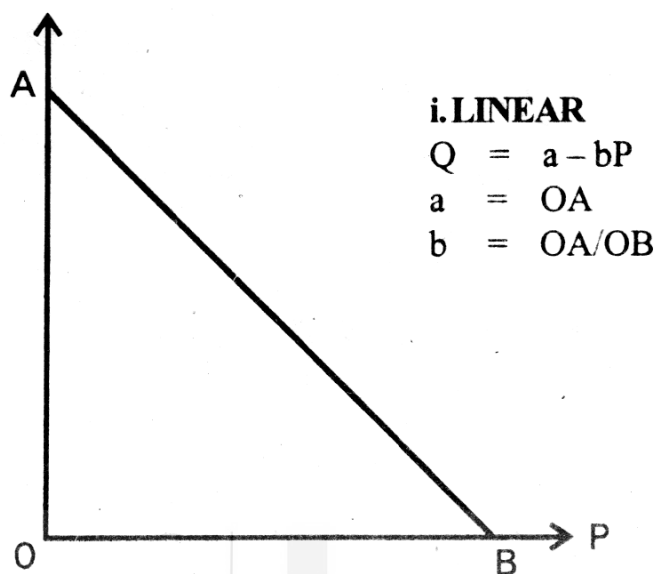


Figure 3.3: Quadratic Function

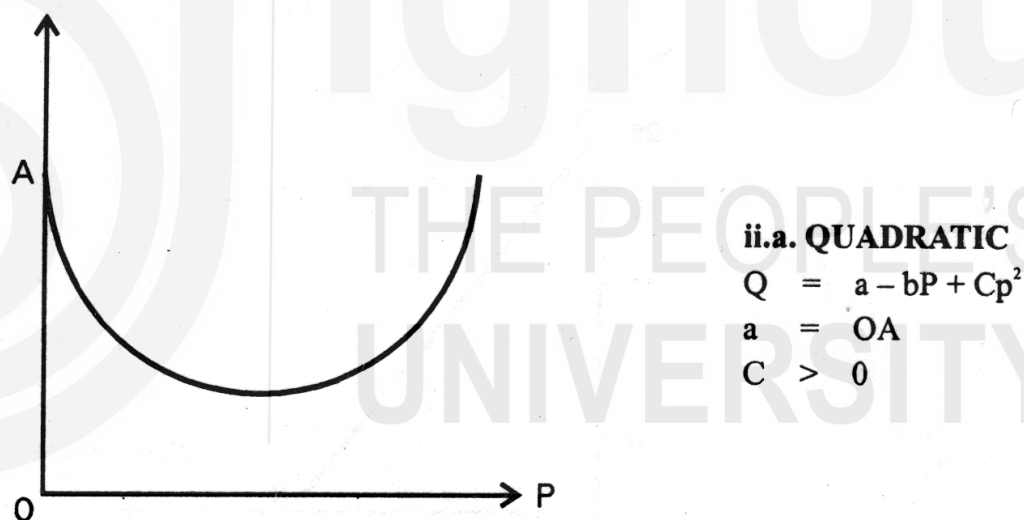


Figure 3.4: Quadratic Function

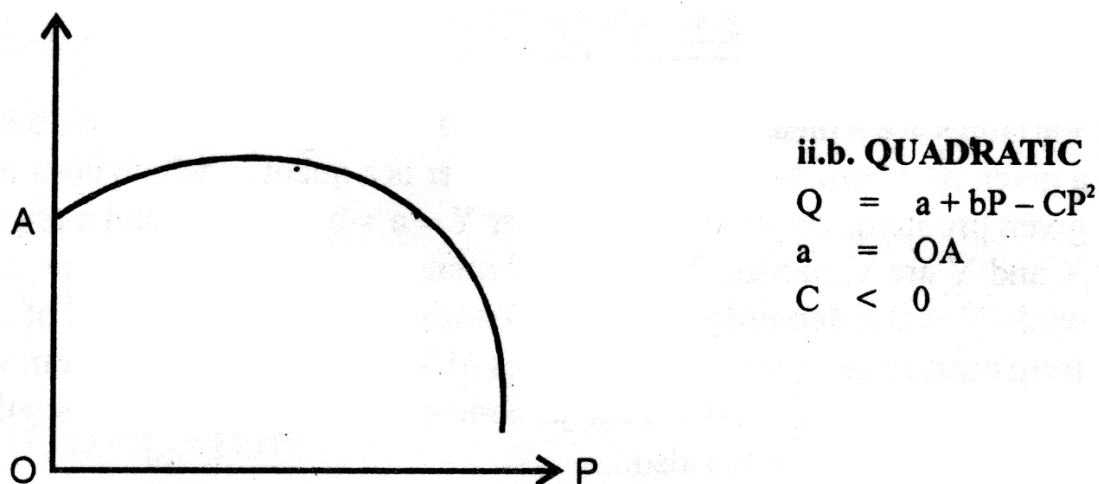
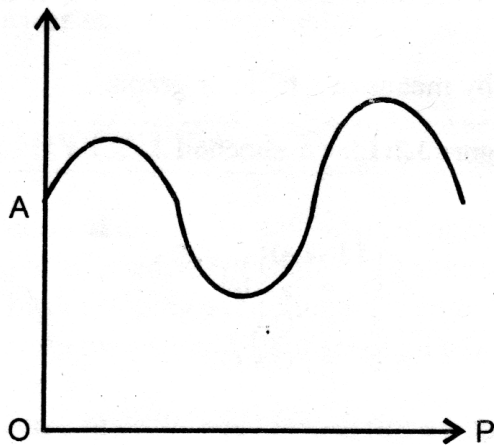


Figure 3.5: Cubic Function



iii. CUBIC

$$Q = a - bP + Cp^2 + dP^3$$

$$a = OA$$

Five key functions of economics are represented graphically:

1. Demand (Linear)

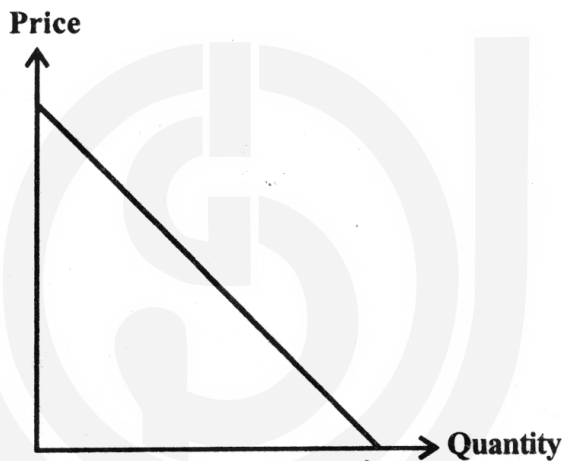


Figure 3.6

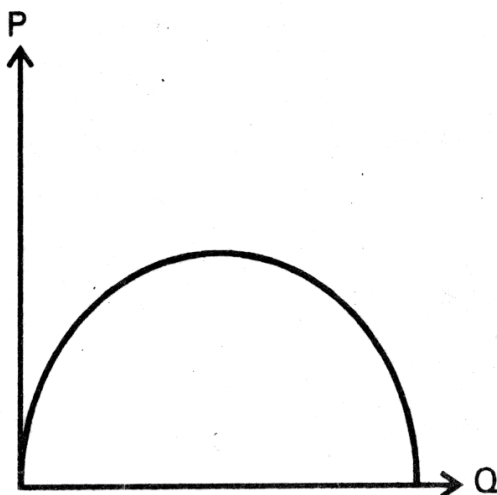
LINEAR

$$Q = a - bP$$

$$P = a + bQ$$

2. Total Revenue (Quadratic)

Figure 3.7

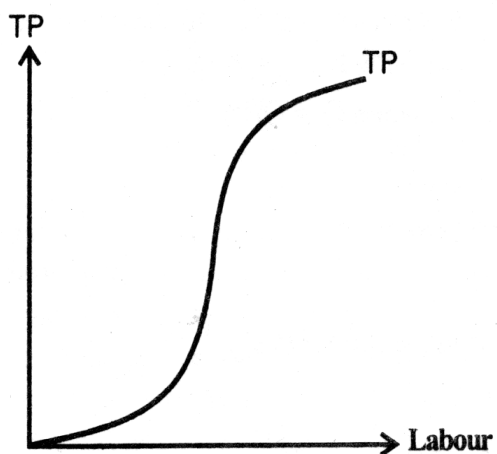


$$TR = a + bQ - CQ^2$$

$$(a = 0)$$

3. Production (Short run) (Cubic)

Figure 3.8

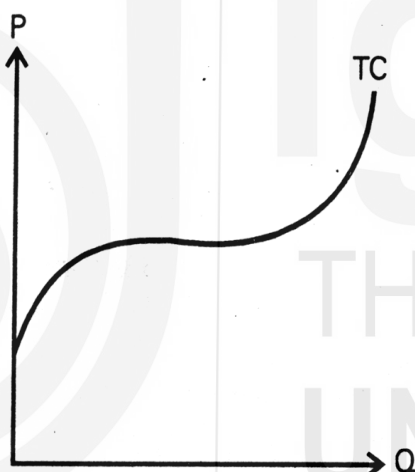


$$TP = a + bL + CL^2 - dL^3$$

(a=0)

4. Cost (Short Run) (Quadratic)

Figure 3.9

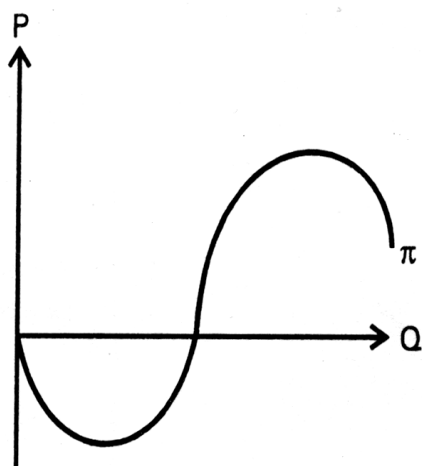


$$TC = a + bQ + CQ^2 - dQ^3$$

(a=0)

5. Profit (Short run) (Quadratic)

Figure 3.10



$$T = a + bQ + CQ^2 + dQ^3$$

(a < 0)

3.4 DERIVATIVES

The “slope” in mathematical use is the concept of ‘marginalism’ in economic use. Thus if $Y=Y(x)$, dy/dx stands for change in Y as result of one unit change in X , i.e. marginal y of x . This slope or marginal function has enormous use in managerial economics. Thus,

$$\frac{dQ}{dP} = \text{Marginal demand of price, when } Q=Q(P)$$

$$\frac{dS}{dA} = \text{Marginal sales of advertisement, when } S=S(A)$$

$$\frac{dR}{dQ} = \text{Marginal revenue of output, when } R=R(Q)$$

$$\frac{dC}{dQ} = \text{Marginal cost of output, when } C=C(Q)$$

In case of ‘averages’

$$\frac{Q}{P} = \text{Average demand}$$

$$\frac{S}{A} = \text{Average sales}$$

$$\frac{R}{Q} = \text{Average revenue}$$

$$\frac{C}{Q} = \text{Average costs}$$

When marginal concept is divided by corresponding average concept, we get the measure of economic concept of elasticity.

$$\frac{dQ}{dP} * \frac{P}{Q} = \text{Price elasticity of demand}$$

$$\frac{dC}{dQ} * \frac{Q}{C} = \text{Output elasticity of cost}$$

$$\frac{dS}{dA} * \frac{A}{S} = \text{Advertisement elasticity of sales revenue}$$

Thus, such elasticities measure the proportion of change, e.g., if the percentage change in demand is greater than the percentage change in price, then

$$\left[\frac{dQ}{dP} * \frac{P}{Q} \right] > 1 \text{ is elastic demand}$$

$$\left[\frac{dQ}{dP} * \frac{P}{Q}\right] < 1 \text{ is inelastic demand}$$

$$\left[\frac{dQ}{dP} * \frac{P}{Q}\right] = 1 \text{ is unitary elastic demand and so on.}$$

Elasticity is discussed in detail in Block 2.

The standard rules of differentiation in calculus are given in Appendix.

3.5 PARTIAL DERIVATIVES

In managerial economics, usually a function of several independent variables is encountered instead of a single variable case shown above. For example, a consumer's demand for a product depends on the price of the product, price of other related goods, income, tastes, etc. When the price changes, the effect on quantity demanded of the goods can be analyzed only when all other variables are kept constant. The functional relationship that is obtained between the quantity demanded of a product and its own price is called a Partial Function (a function of one variable when all other variables are kept constant). The derivatives of the partial functions are known as partial derivatives of the original function and is shown as df/dx_1 or $f_{1(x)}$ or $f'(x)$. The conventional symbol used in maths for the partial derivative is delta, δ . Partial derivatives are functions of all variables entering into the original function $f(x)$.

Examples:

(i) If $Y = f(x_1; x_2)x_1^2x_2^2$

$$\text{Then } \frac{\delta y}{\delta x_1} = 2x_1x_2^2$$

$$\text{And } \frac{\delta y}{\delta x_2} = x_1^2 2x_2 = 2x_2x_1^2$$

(ii) If $Y = \sqrt{x_1x_2} = (x_1x_2)^{1/2} = x_1^{1/2}x_2^{1/2}$

$$= x_1^{1/2}x_2^{1/2}$$

$$\text{Then } \frac{\delta y}{\delta x_1} = \frac{1}{2}x_1^{-1/2}x_2^{1/2} = \frac{1}{2}\frac{x_2^{1/2}}{x_1^{1/2}} = \frac{1}{2}\sqrt{\frac{x_2}{x_1}}$$

$$\text{And } \frac{\delta y}{\delta x_2} = \frac{1}{2}x_1^{1/2}x_2^{-1/2} = \frac{1}{2}\frac{x_1^{1/2}}{x_2^{1/2}} = \frac{1}{2}\sqrt{\frac{x_1}{x_2}}$$

(iii) If $Z = 4x^2 + 3xy + 5y^2$

$$\text{Then } \frac{\delta Z}{\delta y} = 8x + 3y$$

$$\text{Then } \frac{\delta Z}{\delta x} = 3x + 10y$$

Remember the derivatives of a constant is 0, i.e $\delta z/\delta x$ of $5y^2$ is 0. Thus in the above equation while calculating $\delta z/\delta y$ we hold x constant and hence $\delta z/\delta x$ of $4x^2$ is 0. This gives $\delta z/\delta x$ to be $8x+3y$; and $\delta z/\delta y$ to be $3x+10y$.

$\frac{\delta f}{\delta x}$, $\frac{\delta f}{\delta y}$, $\frac{\delta y}{\delta x}$, $\frac{\delta z}{\delta x_2}$, are called first order partial derivatives.

The second order partial derivatives indicate that the function has been differentiated partially twice with respect to a given variable, all other variables being held constant. These are shown (in case of function $Z=f(x,y)$) by $\frac{\delta^2 f}{\delta x^2}$ or f_{xx} and $\frac{\delta^2 f}{\delta y^2}$. Thus $\frac{\delta^2 f}{\delta x^2}$ shows the rate of change of first order partial derivatives f_x with respect to x and y held constant.

Similarly $\frac{\delta^2 f}{\delta y^2}$ is the second order partial derivatives of the function with respect to y with x held constant.

3.6 OPTIMISATION CONCEPT

Optimisation is the act of choosing the best alternative out of the available ones. It describes how decisions or choices among alternatives are taken or should be made. All such optimisation problems have 3 elements:

- Decision Variables:** These are variables whose optimal values have to be determined. For example, a production manager wants to know at what level to set output in order to achieve maximum profit or maximum sales revenue. Here output is the decision or choice variable. Similarly labour, machine, time and raw materials are choice variables if a works manager wants to know what amount of these are to be used so as to produce a given output level at minimum cost. The quantity of any choice variable must be measurable (20kg, 5 labourers, 10 hours, etc.)
- The Objective Function:** It is a mathematical relationship between the choice variables and some variables whose values are to be maximized or minimized. For example, the objective function could relate profit to level of output or cost to amount of labour, machine, time, raw materials, etc. in the above example.
- The Feasible Set:** The available set of alternatives is called a feasible set.

A solution to an optimization problem is that set of values of the choice variables which is in the feasible set and which yields maximum or minimum of the objective function over the feasible set.

Unconstrained Optimization Technique

For unconstrained optimization problem involving single independent variable, we need to satisfy some “conditions”. In economics, the necessary (first order) condition is called the equilibrium condition and sufficient (second order) condition is called the stability condition (continuation of state of equilibrium). There may be equilibrium but it may not be stable. Thus first order condition does not guarantee second order condition. This is summarized in Table 3.1.

Table 3.1: Optimization Conditions

Order	Conditions	Optimization	Unconstrained
First order	Necessary Conditions	Maximization $\frac{\delta y}{\delta x} = 0$	Maximization $\frac{\delta Y}{\delta X} = 0$
Second order	Sufficient Conditions	$\frac{\delta^2 y}{\delta x^2} < 0$	$\frac{\delta^2 y}{\delta x^2} > 0$
	$Y=y(x)$ (assumed)		

Some economic uses of these conditions are discussed below. Given a firm's demand function, $P=45 - 0.5Q$ and the average function, $AC=Q^2-8Q + 57 + 2/Q$, we have to find the level of output Q which

- Maximizes total revenue
- Maximizes profit

Solution

- Since demand function is $P=45-0.5Q$ the total revenue function will be,
 $TR=PQ = (45-0.5Q) Q=45Q - 0.5Q^2$

To maximize TR , we find the derivative and set it to 0 (the first order or necessary condition)

$$\begin{aligned}\text{Now, } \frac{dR}{dQ} &= 45 - 2(0.5)Q \\ &= 45 - Q = 0 \\ \therefore Q &= 45\end{aligned}$$

The second order condition (sufficient conditions) needs $\frac{d^2R}{dQ^2}$ to be negative.

$$\begin{aligned}\text{Since } \frac{dR}{dQ} &= 45 - Q \\ \therefore \frac{d^2R}{dQ^2} &= -1\end{aligned}$$

which is negative. Hence total revenue is maximized when output is 45 units.

b) From profit function

$$\pi = TR - TC$$

$$TC = (AC) \times Q$$

$$= (Q^2 - 8Q + 57 + 2/Q) Q = Q^3 - 8Q^2 - 57Q + 2$$

$$TR = (45 - 0.5Q)Q$$

$$= 45Q - 0.5Q^2$$

After substituting TR and TC, we get

$$\pi = 45Q - 0.5Q^2 - Q^3 + 8Q^2 - 57Q - 2$$

$$\therefore \frac{d\pi}{dQ} = 45 - Q - 3Q^2 - 16Q - 57$$

$$= -3Q^2 + 15Q - 12$$

$$\text{Now set } \frac{d\pi}{dQ} = 0$$

$$= -3Q^2 + 15Q - 12 = 0$$

dividing by -3

$$Q^2 + 5Q + 4 = 0$$

$$\text{or } (Q-4)(Q-1)=0$$

$$\Rightarrow Q = 4 \text{ or } 1$$

Constrained Optimisation Technique

There are many situations where the objective function has to be maximized or minimized subject to certain constraints present in the problem. Thus a consumer may be maximizing utility subject to the income constraint.

The techniques used to analyze such problems are based on that used for unconstrained problems. The constrained problem is converted into unconstrained one with the help of Lagrange Multiplier Technique and then the latter is solved. In this technique, the objective function and constraint is combined in one expression (Lag range expression) such that the constrained maximization or minimization problems are reduced to unconstrained ones.

For example,

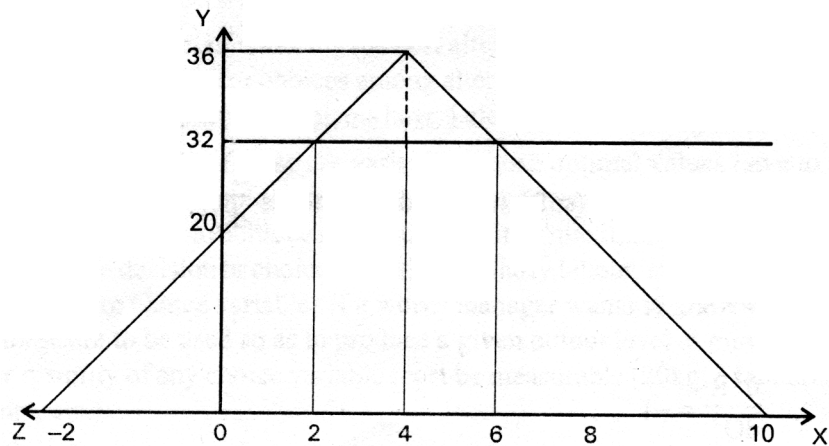
$$\text{Maximize } Y = -x^2 + 8x + 20$$

$$\text{Subject to } x \leq 2$$

(function of single independent variable)

The objective function (function to be optimized) is $Y = -x^2 + 8x + 20$, plotted on graph:

Figure 3.11: Objective Function



The objective function $y = -x^2 + 8x + 20$ can be written as $Y = -(x - 4)^2 + 36$

Now $-(x - 4)^2$ has an unconstrained maximum value of zero at $x = 4$. However, our objective function is $-(x - 4)^2 + 36$, and hence will have an unconstrained maximum of 36 (at $x=4$). This is so for the first derivative test.

The second derivative test $\frac{d^2y}{dx^2}$ gives -2 since $\frac{dy}{dx} = -2x + 8$; $\frac{d^2y}{dx^2} = -2$

Thus both tests give $x=4$ as the above of the variable. This is the value at which the objective function attains unconstrained maximum. However, the problem has a constraint $x=2$. Thus we have to consider the function only up to value of $x=2$ (graph) starting from $x = -2$. The maximum value of the function is then 32 which occurs when $x=2$. Thus the constrained maximum of the function $y = -(x - 4)^2 + 36$ with constraint $x=2$ would occur at $x=2$ and not $x=4$.

The Lagrange expression for constrained maximization is formed as follows $Y = -(x - 4)^2 + 36$ with constraint $x=2$ and $x-2=0$.

Combining, we get the Lagrange expression $L = [-(x - 4)^2 + 36] + \lambda(x - 2)$.

Adding objective function and product of λ (Lagrange Multiplier) and the constraint function $x-2=0$. Now L is a function of x and.

We find

$$\frac{\partial L}{\partial x} \text{ and } \frac{\partial L}{\partial \lambda} \text{ set them to } 0$$

$$\frac{\partial L}{\partial x} = -2(x - 4) + \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = (x - 2) = 0$$

The last equation gives $x=2$. Hence the constrained maximum occurs at $x=2$.

In this problem $\lambda=2(x-4) = 2(2-4) = -4$.

When applying Lagrange technique to solve economic decision problem, λ will have an economic significance. For example in consumer's utility maximizing problem (where quantities of commodities are choice problems), λ will be the marginal utility of money income. In producer's cost minimization problem, λ will be the marginal cost of production. In complex problems, as many λ 's as many there are constraints have to be used.

The problem with two independent variables can also be solved with Lagrangian technique. To find out whether the optimized value is maximum or minimum however, requires second derivative test as well as use of some more determinants not to be discussed here

Activity 1

a) Draw graph of the following functions:

- i) $Q = 10 - 0.4P$
- ii) $Q = 15 - 2P + 4P^2$
- iii) $C = 100 + 0.8Y$

b) A firm's fixed costs are ₹6000.00 regardless of output (they do not change when output changes), variable cost is ₹5 per unit of output (variable cost is dependent on output). Total cost = fixed costs + variable costs. The selling price of the goods is ₹100 per unit. Let Q be the output. State the

- i) Firms fixed cost function
- ii) Variable cost function
- iii) Total cost function
- iv) Total revenue function

c) Maximize $Z = 10xy - 2y^2$
Subject to $x+y=12$

d) Maximize $R = 737 - 8Q^2 + 14A + QA - 4a^2 + 20Q$
Subject to $2Q+A=2$.

Answers:

c) $X=7, y=5, \lambda=-50$

d) $Q=0.52, A=0.96, 1=-6.36$

3.7 REGRESSION ANALYSIS

A manager must often determine the total cost of producing various levels of output. The relation between total cost (C) and quantity (Q) is,

$$C = a + bQ + cQ^2 + dQ^3$$

Where a , b , c and d are parameters of the equation. Parameters are coefficients in an equation that determine the exact mathematical relation among the variables in the equation. If the numerical values of parameters are determined, the manager knows the quantitative relation between output and total cost. If the value of parameters of cost equation are calculated to be $a=1262$, $b=1$, $c=-0.03$ and $d=0.005$, the equation becomes,

$$C = 1262 + 1Q - 0.03Q^2 + 0.005Q^3$$

This equation can be used to compute the total cost of producing various levels of output. If, for example, the manager wishes to produce 30 units of output, the total cost can be calculated as

$$C = 1262 + 30 - 0.03(30)^2 + 0.005(30)^3 = \text{Rs. } 1400.$$

Thus, in order for the cost function to be useful for decision making, the manager must know the numerical value of the parameters.

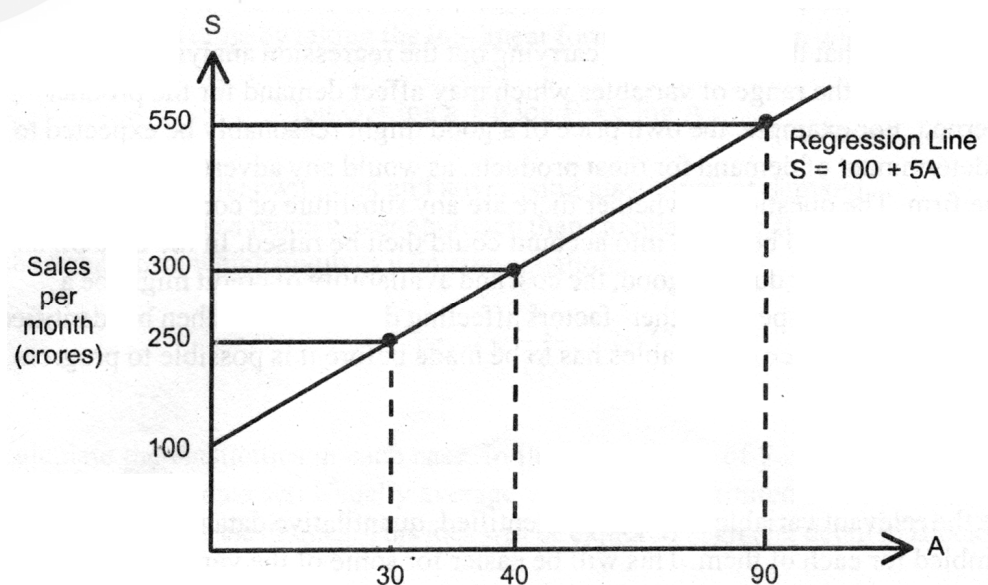
The values of the parameters are often obtained by using a technique called regression analysis. It determines the mathematical relation between a dependent variable and one or more explanatory variables.

- Dependent variable: The variable whose variation is to be explained.
- Explanatory variables: The variables that are thought to cause the dependent variable to take on different values

In the simple regression model, the dependent variable Y is related to only one explanatory variable X , and the relation between X and Y is linear

$$Y = a + bX$$

Figure 3.12: Relation between sales and advertising expenditure



Expenses on advertising per month (crores)

This is the equation for a straight line, with X plotted along the horizontal axis and Y along the vertical axis. The parameter a is called the intercept parameter because it gives the value of Y at the point where the regression line crosses the Y-axis. (X is equal to zero at this point). The parameter b is called the slope parameter because it gives the slope of the regression line. The slope of a line measures the rate of change in Y as X changes (DY/DX); it is therefore the change in Y per unit change in X.

Intercept parameter: The parameter that gives the value of Y at the point where the regression line crosses the Y-axis.

Slope parameter: The slope of the regression line, $b = \Delta Y / \Delta X$, or the change in Y associated with a one-unit change in X.

Y and X are linearly related in a regression model. The effect of change in X on the value of Y is constant. A one-unit change in X causes Y to change by a constant b units.

The figure shows the true relation between sales and advertising expenditures. If a firm chooses to spend nothing on A, its sales are expected to be ₹100 crores per month. If the firm spends ₹30 crores on A then it can expect sales of ₹250 crores ($=100 + 5 \times 30$). $\Delta S / \Delta A = 5$, i.e., for every 1 unit increase on advertising, the firm can expect a 5 unit increase in sales. Regression involves identifying and calculating specific relationships between the independent variables and the dependent variable. It involves a number of stages which are described in another section.

Activity 2

- The slope parameter is and the intercept parameter is in the equation $R=a+bW$.
- In Figure 3.12, what will be the monthly sales of the firm, if the advertising expenditure is increased from ₹30 crores to ₹40 crores per month?

.....
.....
.....
.....

3.8 SPECIFYING THE REGRESSION EQUATION

The first thing that the organisation carrying out the regression analysis needs to do is to determine the range of variables which may affect demand for the product concerned. For example, the own price of a good might reasonably be expected to be a determinant of demand for most products, as would any advertising being done by the firm. The question of whether there are any substitute or complementary goods which need to be taken into account could then be raised. In the case of, an expensive consumer durable good, the cost and availability of credit might be a consideration. Any special 'other' factors affecting demand could then be identified and so on. This choice of variables has to be made before it is possible to progress to the next stage.

Data Collection

Once the relevant variables have been identified, quantitative data need to be assembled for each of them. This will be easier for some of the variables than for others. In dealing with an established product, for example, the firm might reasonably be expected to have access to a range of information regarding the variables which it controls such as own price and advertising. What may be more difficult to obtain, however, is information about competitors' products. On this front, price data can be obtained through observing retail prices, as this information by definition is in the public domain and cannot be hidden. This requires continued market observation, perhaps over a long period of time. Likewise, information about product design changes can be obtained by buying the competitors' product(s), but this may be expensive if there are many on the market. Confidential, commercially sensitive information such as actual advertising expenditure by competitors and their proposed new products present much more difficult problems in terms of access and may have to be left out of the process altogether. Data on levels of disposable income, population variables, interest rates and credit availability are easier to obtain, for example from government statistics, but other variables are more problematic. How can things like expectations and tastes be measured for instance? In these cases the available data, perhaps resulting from market surveys, may be qualitative rather than quantitative. Some means of conversion need to be found if they are to be included in the regression analysis at all. These are the things which the decision maker needs to keep in mind while collecting and selecting data on the relevant variables. Once the first two steps have been completed, the next stage is to specify the likely form of the regression equation. There are two main forms which are used in practice -the linear demand function and the non-linear or power function. Both treat the demand for the product as the dependent variable, while the independent variables are those which have previously been identified as having an effect on demand. If, for example, the firm had decided that the only variables affecting demand for a particular product with its own price and advertising levels then the linear demand function would be written as:

$$Q = a + bP + cA$$

Alternatively, under these conditions the exponential (power) function would be written as:

$$Q = a P^b A^c$$

In each case, the 'a' term represents the intercept of the-line drawn from the equation with the vertical axis. The 'b' and 'c' terms represent the regression coefficients with respect to own price and advertising respectively. These show the impact of each of these variables on product demand. Once they have been estimated it is possible to predict the level of demand for any set of values of the independent variables simply by substituting them into the equation.

The exponential form of the equation has the advantage that it can be rewritten to give direct estimates of the respective elasticities of demand for the independent variables. This is done by taking the log-linear form of the equation which in this case would be:

$$\log Q = \log a + b \log P + c \log A$$

Where 'b' and 'c' are the own price and advertising elasticities of demand respectively. This is a much easier approach than calculating elasticities through use of the linear form which involves using the equation:

$$ED = b \frac{P_D}{Q_D}$$

to calculate the elasticities in each case. In this case values of P and Q need to be obtained from the data set. Usually average values are substituted in the above equation to estimate elasticities. This idea will be explored in greater detail in Block 2.

Which of the two forms of equation is chosen depends upon the expected relationship between the variables being included. In practice, however, the actual relationship between them may not be known in advance. In this case, the decision maker may experiment with both forms of equation in order to find the one which most closely fits the data.

Activity 3

1. What are the things which the decision maker needs to keep in mind while collecting and selecting data on the relevant variables?

.....

.....

.....

.....

2. Show that estimated coefficients using a log-linear technique are estimates of elasticity with respect to the relevant variable.

.....

.....

.....

.....

3.9 ESTIMATING THE REGRESSION EQUATION

Consider a firm with a fixed capital stock that has been rented under a long-term lease for ₹ 100 per production period. The other input in the firm's production process is labour, which can be increased or decreased quickly depending on the firm's needs. In this case, the cost of the capital input (₹100) is fixed and the cost of labour is variable. The manager of the firm

wants to know the relationship between output and cost, that is, the firm's total cost function. This would allow the manager to predict the cost of any specified rate of output for the next production period.

Specifically, the manager is interested in estimating the coefficients 'a' and 'b' of the function

$$Y = a + bX$$

where the dependent variable Y is total cost and the independent variable X is total output. If this function is plotted on a graph, the parameter 'a' would be the vertical intercept (i.e., the point where the function intersects the vertical axis) and 'b' would be the slope of the function. Recall that the slope of a total function is the marginal function. As $Y = a + bX$ is the total cost function, the slope, 'b', is marginal cost or the change in total cost per unit change in output.

Assume that data on cost and output have been collected for each of seven production periods and are reported in Table 3.2. Note that there is a cost of Rs. 100 associated with an output rate of zero. This represents the fixed cost of the capital input, which must be paid regardless of the rate of output. These data are shown as points in Figure 3.1. They suggest a definite upward trend, but they do not trace out a straight line. The problem is to determine the line that best represents the overall relationship between Y and X. One approach would simply be to "eyeball" a line through these data in a way that the data points were about equally spaced on both sides of the line. The coefficient 'a' would be found by extending that line to the vertical axis and reading the Y-coordinate at that point. The slope, 'b', would be found by taking any two points on the line, $\{X_1, Y_1\}$ and $\{X_2, Y_2\}$ and using the slope formula

$$b = \frac{Y_2 - Y_1}{X_2 - X_1}$$

Although this approach could be used, the method is quite imprecise and can be employed only when there is just one independent variable. What if production cost depends on both the rate of output and the size of the plant? To plot the data for these three variables (total cost, output, and plant size) would require a three dimensional diagram; it would be nearly impossible to eyeball the relationship in this case. The addition of another independent variable, say average skill levels of the employees, would place the data set in the fourth dimension, where any graphic approach is hopeless.

Table 3.2: Hypothetical Data on Total Cost and Total Output

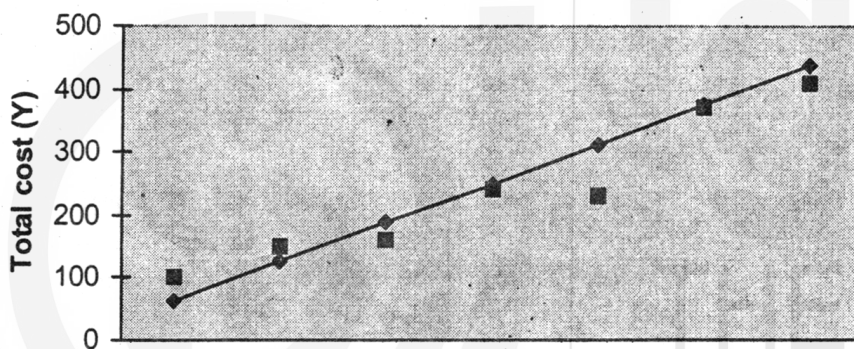
Production period	Total Cost (Y_1) ₹	Total Output (X_1)
	100	0
	150	5
	160	8
	240	10
	280	15
	370	23
	410	25

There is a better way. Statisticians have demonstrated that the best estimate of the coefficients of a linear function is to fit the line through the data points so that the sum of squared vertical distances from each point to the line is minimized. This technique is called ordinary least-squares regression (OLS) estimation and a number of statistical packages, including excel. Based on the output and cost data in Table 3.1, the least-squares regression equation will be shown to be

$$\hat{Y} = 87.08 + 12.21X$$

This equation is plotted in Figure 3.13. Note that the data points fall about equally on both sides of the line. Consider an output rate of 5. As shown in Table 3.2, the actual cost associated with this output level is 150. The value predicted by the regression equation, referred to as $\hat{Y}$, is 148.13. That is, $\hat{Y} = 87.08 + 12.21(5) = 148.13$. The deviation of the actual Y value from the predicted value (i.e., the vertical distance of the point from the line), $\hat{Y}_1 - \hat{Y}$ is referred to as the residual or the prediction error.

Figure 3.13: Total Cost, Total Output and Estimated Regression Equation



There are many values that might be selected as estimators of 'a' and 'b', but only one of those sets, defines a line that minimizes the sum of squared deviations [i.e., that minimizes $\sum(Y_i - \hat{Y}_i)^2$]. The equations for computing the least-squares estimators and $\hat{b}$ are

$$\hat{b} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$$

and

$$\hat{a} = \bar{Y} - \hat{b}\bar{X}$$

When $\bar{Y}$ and $\bar{X}$ are the means of the Y and X variables.

Using the basic cost of output data and the example, the necessary calculations are shown in Table 3.3. Substituting the appropriate values into the following equations, the estimates $\hat{a}$ of $\hat{b}$ and are computed to be

$$\hat{b} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} = 12.21$$

$$\hat{a} = \bar{Y} - \hat{b}\bar{X} = 237.14 - 12.21(12.29) = 87.08$$

Table 3.3: Summary calculation for computing the Estimates $\hat{a}$ and $\hat{b}$

Cost (Y_i)	Output (X_i)	$Y_i - \bar{Y}$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$
100	0	-137.14	-12.29	151.04	1,685.45
150	5	-87.14	-7.29	53.14	635.25
160	8	-77.14	-4.29	18.40	330.93
240	10	-002.86	-2.29	5.24	-6.55
230	15	-7.14	2.71	7.34	-19.35
370	23	132.86	10.71	114.70	1,422.93
410	25	172.86	12.71	161.54	2,197.05
$\bar{Y} = 237.14$	$\bar{X} = 12.29$			$\Sigma(X_i - \bar{X})^2$ =511.40	$\Sigma(X_i - \bar{X})(Y_i - \bar{Y})$ =6,245.71

Thus the estimated equation for the total cost function is

$$\hat{Y} = 87.08 + 12.2X$$

The estimate of the coefficient a is 87.08. This is the vertical intercept of the regression line. In the context of this example, $\hat{a} = 87.08$ is an estimate of fixed cost. Note that this estimate is subject to error because it is known that the actual fixed cost is ₹ 100. The value of $\hat{b}$ is an estimate of the change in total cost for a one-unit change in output (i.e., marginal cost). The value of $\hat{b}$, ₹ 12.21, means that, on an average, a one-unit change in output results in ₹ 12.21 change in total cost. Thus $\hat{b}$ is an estimate of marginal cost.

Activity 4

Suppose, for example, that the estimation process had given the following figures for the coefficients:

$$\log QD = \log 200 - 1.5 \log P + 2.4 \log A$$

where QD is quantity demanded

P is own price and

A is advertising expenditure

What can we deduce from the estimated equation?

.....

.....

.....

.....

3.10 DECISION UNDER RISK

The focus of this section is decision making under risk. The objective will be to develop guidelines for making rational decisions given the decision makers attitudes towards risk. Attitudes towards risk may be of three types: (a) A *risk-seeker* is one who prefers risk, given a choice between more or less risky investments, with identical expected money returns; he will select the riskier investment. (b) A *risk averter* is one who faced with the same situation will select the less risky investment. (c) A *risk-neutral* person is one who faced with the same situation will be indifferent to the choice. For him any investment is equally preferable to the other. It is difficult to slot people in one of these categories. You would perhaps have observed both risk averse and risk seeking behaviour in the real world.

The analysis of risk is based largely on the concepts of probability and probability distribution that are commonly encountered in elementary statistics. First the terms strategy, states of nature and outcomes need to be defined. A strategy is one of the many alternative plans or courses of action that could be implemented in order to achieve managerial goal. A manager might be considering three strategies to increase profits - build a more modern plant which may produce at low cost, implement a new marketing programme to increase sales or change the design of product to decrease cost and increase sales.

A state of nature is a condition that may exist in the future and that will have a significant effect on the success of a strategy. For example, the manager may not be aware of the economic conditions in the future. The possible states of nature may be normal, recession or boom.

The outcome results in either gain or loss based on a particular combination of strategy and state of nature. The decision maker has no control over the states of nature that will prevail in future but the future states of nature will certainly affect the outcome of any strategy that he or she may adopt. The particular decision made will depend, therefore, on the decision maker's knowledge or estimation of how a particular future state of nature will affect the outcome of each particular strategy.

Given a set of outcomes, X_i , and the probability of each occurring, P_i , three statistics relating to probability distributions can be used.

- The expected value of mean is a measure of expected return. This is represented by

$$\mu = \sum_{i=1}^n P_i(X_i)$$

Where P_i is the probability of X_i is the outcome.

- The standard deviation is a measure of risk. This is represented by

$$\sigma = \sqrt{\sum_{i=1}^n P_i (X_i - \mu)^2}$$

- The coefficient of variation is a measure of risk per unit of money of return.

$$v = \frac{\sigma}{\mu}$$

These statistics have a direct application in measuring the expected return and risk associated with any business decision for which a set of outcomes and their probabilities have been determined. The expected value, standard deviation, and the coefficient will be referred to as risk-return evaluation statistics.

Having defined risk and reviewed some of the related terminology, the task now is to develop quantitative measures of return and risk and to show how they are applied in decision making. We know that individuals have different preferences concerning risk taking. It is also important that such preferences be identified and their effect on decisions evaluated. Rational decision making requires that the expected return be determined and the risk be measured, and that there be information about the manager's preference for risk. The expected value, the standard deviation, and the coefficient of variation will be referred to as risk-return evaluation statistics.

Let us take an example where two investments, I and II, are being considered. Both investments require an initial cash outlay of Rs.100 and have a life of five years. The return on each depends on the rate of inflation over the five-year period. Of course, the inflation rate is not known with certainty, but suppose that the collective judgment of economists is that the probability of no inflation is 0.20, the probability of moderate inflation is 0.50, and the probability of rapid inflation is 0.30. The outcomes are defined as the present value of net profits for the next five years. These outcomes for each state of nature (i.e., rate of inflation) for each investment are shown in table 3.3.

Analysis of these investments can be made by calculating and comparing the three evaluation statistics for each alternative. The expected value μ is an estimate of the expected return for the investment. Because risk has been defined in terms of the variability in outcomes, the standard deviation s is a measure of risk associated with the investment. The larger the μ the greater is the risk. Risk per rupee of expected return is measured by the coefficient of variation (v).

The evaluation statistics for each investment alternative are computed as follows:

$$\mu_i = \sum_{i=1}^n P_i(X_i) = 0.2(100) + 0.5(200) + 0.3(400) = 240$$

$$\sigma = \sqrt{\sum_{i=1}^n P_i (X_i - \mu)^2} = \sqrt{0.2(100 - 240)^2 + 0.5(200 - 240)^2 + 0.3(400 - 240)^2}$$

$$= 111.36$$

$$\vartheta_I = \frac{\sigma_I}{\mu_I} = 111.36/240 = 0.46$$

$$\mu_{II} = 0.2(150) + 0.5(200) + 0.3(250) = 205$$

$$\sigma_{II} = \sqrt{0.2(150 - 205)^2 + 0.5(200 - 205)^2 + 0.3(250 - 205)^2} = 0.35$$

$$\vartheta_{II} = \frac{0.35}{205} = 0.17$$

The expected return for investment I of ₹ 240 is higher than for II ₹.205, but I is a riskier investment because $\sigma_I = 111.36$ is greater than $\sigma_{II} = 0.35$. Also, risk per dollar of expected returns for I ($\vartheta_I = 0.46$) is higher than for ($\vartheta_{II} = 0.17$). Which is the better investment? The choice is not clear. It depends on the investor's attitude about taking risks. A young entrepreneur may well prefer I, whereas an older worker investing in a retirement account where risk ought to be minimized might prefer II. Generally higher returns are associated with higher risk.

Table 3.3: Probability Distribution for Two Investment Alternatives

State of Nature	Probability (P _i)	Outcome(X _i)
Investment I		
No inflation	0.20	100
Moderate inflation	0.50	200
Rapid inflation	0.30	400
Investment II		
No inflation	0.20	150
Moderate inflation	0.50	200
Rapid inflation	0.30	250

Decision Tree

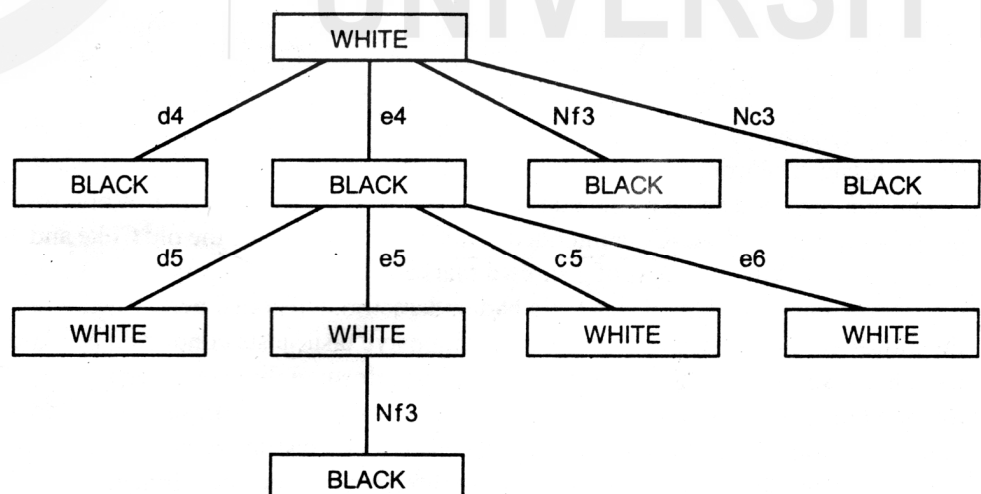
Some strategic decisions are based on a sequence of decisions, states of nature and possibly even strategic decisions. Alternative strategies can be evaluated then, by using a decision tree, which traces sequences of strategies and states of nature to arrive at a set of outcomes. The probability of each outcome is found by multiplying the probabilities on each branch leading to that outcome.

A decision tree shows two or more branches at each point where a decision or event (state of nature) leads to the various outcomes. The decision tree approach can be directly applied to managerial decision-making. A firm entering a new market may decide to build a small or large plant (managerial decisions). This has no probabilities. But there may be stochastic elements (an outcome determined by chance) associated with each decision e.g., reaction of a major competitor and the economic condition. The competitor may react by starting a national, regional, or a new advertising program. The probability of each occurrence will depend on the size of the plant.

The possible economic conditions and then probabilities may be recession, normal or boom. Here also the probability will depend on the size of the plant. The probability of each combination is found by multiplying the probability along each of the branches leading to the outcome. For example, if the manager decides to build a large plant there is a 70 per cent chance that the competitor will respond with a national advertisement. We are given that there is a 30 per cent chance of a boom. The probability associated with the outcome of 80 is therefore 0.21 ($=0.7 \times 0.3$). Similar for other entries in the decision tree.

Decision trees are particularly useful if sequential decision-making is involved. In a game of chess, white has the first move. White has several options at this stage. To keep the problem tractable, let us assume that there are four possible moves for white : (i) move the king's pawn two squares; (ii) move the queen's pawn two squares; (iii) move the king's knight to king bishop three; and (iv) move the queen's knight to queen bishop three. In chess notation, the four moves are - (i) e4; (ii) d4; (iii) Nf3; and (iv) Nc3. Once white has made the first move, Black has several different moves at his disposal. To keep the problem tractable, let us follow the decision tree only when white has moved e4. Even then, Black has several moves at his disposal. Let us assume that he has only four options - (i) move the king's pawn two squares (e5), (ii) move the queen's pawn two squares (d5); (iii) move the queen's bishop's pawn two squares (c5); and (iv) move the king's pawn one square (e6). Once white has moved e4 on the first move and Black has moved e5 on the first move, white has several moves at his disposal. One of them is, move the king's knight to bishop three (Nf3). And so the game goes on.

Figure 3.17 : Decision Tree



Each vertex or node indicates a decision to be taken by one of the players, the number within the rectangle indicating whose turn it is to move. There need not actually be two players, one of the players can be regarded as nature or chance. The main advantage of using a decision tree is that it helps one to isolate each chain and follow it through to the very end.

3.11 UNCERTAINTY ANALYSIS AND DECISION MAKING

Certainty appears to be a theoretical and impractical state, as here the investor has perfect knowledge of the investment environment such that he is definite about the size, regularity and periodicity of flow of returns. Such situations may exist in the short-run (e.g. fixed deposit in a nationalised bank). However, long-run or long-range investments are not predictable as they are influenced by many kinds of changes taking place with time: political, economic, market and technology etc.

Risk is more common in the real world. A situation with more than one possible outcome to a decision such that the probability of each of these outcomes can be measured is a risk situation. For example, tossing of a coin (i.e. 50-50) or investing in a stock. The greater the number and range of outcomes, the greater is the risk associated with the decision or action. Uncertainty is a situation where there is more than one possible outcome to a decision but the probability of each specific outcome occurring is not known or even meaningful. This may be due to insufficient information or instability in the nature of variables. In extremes cases of uncertainty, the outcomes may itself be not clear. Decision making under uncertainty is necessarily subjective.

A Case Study of an American multinational chain of fast food restaurants Burger's which failed miserably.

This is a case study of a well-known American fast-food company which offers affordable as well as high quality products. On an average 11 million people visit the restaurants all over the world. Burger's are one of the most important products of fast food companies; The Fast food restaurant itself gets one order of burger for every two orders.

When health trends set in, the fast food restaurant noticed that its burger sales were dipping in USA. To attract health-conscious customers back it decided to launch new type of healthier burgers which contained 40% less fat and 30% less calories. They launched new type of burgers based on several assumptions such as their customers wanted healthier options and were ready to pay more for that. It was a very calculated risk taken by the company based on the ongoing trend of health.

The product which was launched as a game changer in fast food market, failed to attract the customers and it was discontinued within one year. It was a major blow to the global fast food giant which was struggling to provide fast food at its restaurants. There were several explanations given by experts on why this product could not stabilize. But one thing is clear, what customers say/want and do are completely different things altogether, so there is always uncertainty in business environment which cannot be eliminated.

This case clearly shows that even a well-conceived strategy is risky and can lead to results estimated to have a small probability of occurrence. Indeed the failure rate for new products in the United States is a stunning 80 percent.

3.12 ROLE OF MANAGERIAL ECONOMIST

In general, managerial economics can be used by the goal-oriented manager in two ways. First, given an existing economic environment, the principles of managerial economics provide a framework for evaluating whether resources are being allocated efficiently within a firm. For example, economics can help the manager if profit could be increased by reallocating labour from a marketing activity to the production line. Second, these principles help the manager to respond to various economic signals. For example, given an increase in the price of output or the development of a new lower-cost production technology, the appropriate managerial response would be to increase output. Alternatively, an increase in the price of one input, say labour, may be a signal to substitute other inputs, such as capital, for labour in the production process.

Thus, the working knowledge of the principles of managerial economics can increase the value of both the firm and the manager.

3.13 SUMMARY

Various quantitative tools are used by the manager to help him in making decisions.

An opportunity set is a set of alternative actions which are feasible. Variables are things which can change and can take a set of possible values within a given problem. A function shows the relation between two variables. It can take different forms-linear, quadratic, cubic. Partial derivatives are functions of all variables entering into the original function $f(x)$. Optimisation is the act of choosing the best alternative out of all available ones. Regression analysis helps to determine values of the parameters of a function - Economic analysis of risk becomes crucial with reference to decisions.

3.14 SELF-ASSESSMENT QUESTIONS

1. Find the present value of Rs.10,000 due in one year if the discount rate is 5 per cent, 8 per cent, 10 per cent, 15 per cent, 20 per cent and 25 per cent.
2. Apply the decision making model to your decision to attend college at MBA level.
3. Discuss with examples how managerial economics is an integral part of business activity.
4. Suppose a seller has two markets to serve. The demand schedules in them are given in the table. Suppose that he has 1400 units to sell and maximize profits thereby. What prices will he set in the two markets? Apply equi-incrementalism principle to get your answer. Could you have applied equi-marginalism.

Market A		Market B	
Price ₹	Quantity	Price ₹	Quantity
50	400	60	600
40	600	50	800
30	900	40	1100
20	1000	34	1400

[Hint: First get total revenue in each market by multiplying price with quantity.]

- A firm is producing two products x and y , and has the following profit function $P = 64x - 2x^2 + 4xy - 4y^2 + 32y - 14$. Find the profit Maximizing levels of output for each of the two products. (Ans.: $x = 40$, $y = 24$, $p = 1650$).
- Maximize $Z = 10xy - 2y^2$
Subject to $x + y = 12$
- What are central or basic problems of an economy?
- Which problems of an economy constitute the subject matter of microeconomics

3.15 FURTHER READINGS

Koutsoyiannis, A. (2018). *Modern Microeconomics* (2nd Ed.). Macmillian Publishers India

Baumol, W.J.(1979). *Economic Theory and Operations Analysis* (4th Ed.) Prentice Hall India, New Delhi.

Lewis, W.C., Jain, S.K., & Peterson, H.C. (2005). *Managerial Economics* (4th Ed.). Pearson.

The standard rules of differentiation in calculus are given below:

Appendix

- A. Basic rule: $Y = ax_n$
$$\frac{dy}{dx} = naX^{n-1}$$
- B. Additional rule: $Y = u(x) + v(x)$
$$\frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx}$$
- C. Product rule: $Y = u(x) * v(x)$
$$\frac{dy}{dx} = u(x) \frac{dv}{dx} + v(x) \frac{du}{dx}$$
- D. Quotient rule: $Y = \frac{u(x)}{v(x)}$
$$\frac{dy}{dx} = \frac{v(x) \frac{du}{dx} - u(x) \frac{dv}{dx}}{v^2}$$
- E. Chain rule: $Y = y[u(x)]$
$$\frac{dy}{dx} = \frac{dy}{du} * \frac{du}{dx}$$
- F. Logarithm rule: $Y = \log X$
$$\frac{dy}{dx} = \frac{1}{x}$$
- G. Exponential rule: $Y = ex$
$$\frac{dy}{dx} = e^x$$

Block 2

DEMAND AND REVENUE ANALYSIS

Unit 4: Demand Concepts and Analysis	73
Unit 5: Demand Elasticity	89
Unit 6: Demand Estimation and Forecasting	108

BLOCK 2 DEMAND AND REVENUE ANALYSIS

This block examines demand, elasticity and forecasting techniques. Demand is the force that drives all business. Without a demand for its goods or services, a firm is doomed to failure. Several determinants of demand are identified in **UNIT 4**. The most important of these are the product's price, the level of income, consumers' tastes and preferences, and the prices of complementary or substitute goods. When drawing a demand curve, only the axis of price and quantity demanded are used; all determinants of demand other than price are assumed to be constant. Although demand curves allow us to examine the relationship between price and quantity demanded, the analysis is a simplification of conditions in the real world.

Unit 5 goes a step further and discusses the importance of elasticities as measures of the responsiveness of one item to changes in another item. The price elasticity of demand measures the responsiveness of quantity demanded to changes in price (while holding all other things constant). Likewise, income elasticity measures the responsiveness of quantity demanded to changes in income (again holding all other things constant). Cross-price elasticity of demand measures the responsiveness of quantity demanded to changes in the price of another good.

These elasticities are calculated using two different techniques – arc elasticity and point elasticity. The arc elasticity method of calculation estimates each of these three elasticities by using past data on changes in quantity demanded and the changes in the particular determinant selected (price, income, or the price of another goods). The second method of estimation, called the point estimation method, requires an accurate estimate of the demand for its calculation. The point estimate is more precise, but is also requires the estimator to have more information to carry out the calculation.

In order to actually estimate and forecast real world demand, we use a variety of methods including regression analysis in **Unit 6**. By using regression, all the determinants of a particular demand can be included in the analysis at once. Use of regression estimates also allows a comparison of the importance and magnitude of the various determinants of demand and the easy calculation of elasticities.

You should appreciate that the need to estimate the forecast demand is not confined to private sector organizations. It is equally necessary for many public sector organizations. The principles of demand forecasting for public sector organizations are not significantly different from those applied in the private sector. However, this block focuses on the range of demand forecasting techniques often used in the private sector.

UNIT 4 DEMAND CONCEPTS AND ANALYSIS

Objectives

After studying this unit, you should be able to:

- **define** demand and its determinants;
- **explain** the Law of Demand;
- **identify** differences between Firm's Demand Curve and Market Demand Curve.

Structure

- 4.1 Introduction
- 4.2 The Demand Function
- 4.3 The Law of Demand
- 4.4 The Market Demand Curve
- 4.5 The Determinants of Demand
- 4.6 The Product's Price as a Determinant of Demand
- 4.7 Income as a Determinant of Demand
- 4.8 Tastes and Preferences as Determinants of Demand
- 4.9 Other Prices as Determinants of Demand
- 4.10 Other Determinants of Demand
- 4.11 Summary
- 4.12 Key Words
- 4.13 Self-Assessment Questions
- 4.14 Further Readings

4.1 INTRODUCTION

Demand refers to the quantities of goods that consumers are *willing and able* to purchase at various prices during a given period of time. For your demand to be meaningful in the marketplace you must be *able* to make a purchase; that is, you must have enough money to make the purchase. There are, no doubt, many items for which you have a willingness to purchase, but you may not have an effective demand for them because you don't have the money to actually make the purchase. For example, you might like to have a 3600 -square-foot resort in Mussorie, an equally large beach house in Goa, and a private jet to travel between these places on weekends and between semesters. But it is likely that you have a budget constraint that prevents you from having these items.

For demand to be effective, a consumer must also be willing to make the purchase. There are many products that you could afford (that is, you have the ability to buy them), but for which you may not be willing to spend your income. Each of us has a unique perspective on our own personal satisfaction and the things that may enhance that satisfaction. The important point is that if you do not expect the consumption of something to bring you added satisfaction, you will not be willing to purchase that good or service.

Therefore, you do not have a demand for such things despite the fact that you might be able to afford them.

When we discuss demand, we are always referring to purchases made during a *given period of time*. For example, you might have a weekly demand for soft drinks. If you are willing and able to buy four soft drinks at a price of Rs 5.00 each, your demand is four soft drinks a week. But your demand for shoes may be better described on a yearly basis so that, at an average price of Rs. 800.00 a pair, you might buy three pairs of shoes per year. The important point here is that when we refer to a person's demand for a product, we usually mean the demand over some appropriate time period, not necessarily over the rest of the person's life.

Think about the last time you spent money. It could have been spent on a car, a computer, a new tennis racquet, or a ticket to a movie, among literally thousands of other things. No matter what you purchased, you decided to buy something because it would please you. You are not forced to make purchases. You do so because you expect them to increase your personal satisfaction.

If these things give us satisfaction, we say that they have value to us. Used in this way, value implies value in use. Air has a value in use, because we benefit from breathing air. But air is free. If air has value to us, why is it free? We certainly would be willing to pay for air rather than do without it. But air is available in such abundance that we treat it as a free good. We also get satisfaction from using petrol. Petrol has value in use. But unlike air, we must pay for the petrol we use. That is, petrol has value in exchange as well as value in use. We are willing to exchange something-usually money-for the use of some petrol.

Why is air free, but petrol is costly? One important reason is that petrol is scarce, whereas air is abundant. This should start making you think about the role that scarcity plays in the economy. But be careful as you do so. Just because something is scarce does not necessarily mean it will have value in exchange. Another reason that something may not have value in exchange is because it has no value in use. That is, people just do not get any satisfaction from possessing or using it.

We all have a limited amount of money that we can exchange for goods and services. The limit varies from individual to individual. For example, a school teacher typically has far less money to spend than a successful investment banker. An unskilled labourer has less money to exchange for goods and services than a skilled labourer. However, we all (even the richest among us) have a limited amount of money for buying things that can bring us satisfaction. As a result, we all make decisions about how we will spend, save, and/or borrow money. This implies that how we choose to allocate our money is an important factor in determining the demand for various goods and services in the economy.

The demand function sets out the variables, which are believed to have an influence on the demand for a particular product. The demand for different products may be determined by a range of factors, which are not always the same for each of them. The presentation in this section is of a generic demand function which includes some of the most common variables that affect demand. For any individual product, however, some of these may not apply. Thus, any attempt by the firm to predict demand for a product on the basis of the demand function will require some initial knowledge, or at least informed guesswork, about the likely influences on it.

The demand function can be written as:

$$Q_d = f(P_o, P_c, P_s, Y_d, T, A, CR, R, E, N, 0)$$

The first three variables in the function relate to price. They are the own price of the product (P_o), the price of complements (P_c) and the price of substitutes (P_s) respectively. In the case of the own price of a good, the expected relationship would be, *the higher the price the lower the demand*, and *the lower the price the higher the demand*. This is the law of demand which is explained in greater detail in the next section. In the case of complements, if the *price of complementary goods* increases, we would expect demand to fall both for it and for the good that it is complementary to. This is the case as fewer people would now wish to buy either good given that the complementary good is now more expensive and this has the effect of reducing demand for the other good as well. In contrast, if the price of a substitute good rises, then demand for the good that it is a substitute for would be expected to rise as people switched to buying the latter rather than its more expensive substitute. Complements and substitutes are also explained in detail later on.

The fourth variable is the demand function, Y_d stands for disposable income, that is, the amount of money available to people to spend. The greater the level of disposable income, the more people can afford to buy and hence the higher the level of demand for most products will be. This assumes of course that they are 'normal' goods, purchases of which increase with rising levels of income, as opposed to 'inferior' goods that are purchased less frequently as income rises. The use of disposable income rather than just income is justified on the grounds that people do not have total control over their gross incomes. There will, for example, be deductions to be made in the form of taxes. Thus, the level of disposable income can change over time, for example changes in tax rates.

The effect of changes in disposable income on the demand for individual products will of course be determined by the ways in which it is spent. This is where the fifth variable, tastes (T), needs to be taken into account. Over a period of time, tastes may change significantly, but this may incorporate a wide range of factors. For example, in case of food, greater availability of alternatives may have a significant effect in changing the national diet. Thus, in India for instance, the demand for bajra has fallen over the past 10 years as people have switched to eating rice and wheat instead. Social pressures may also act to alter tastes and hence demand. For example, tobacco companies

have been forced to seek new markets as smoking has become less socially acceptable in the USA and Western Europe, thus reducing demand in these areas. Changes in technology may also have an impact. For example, as the demand for smart televisions increased, the demand for normal televisions fell as tastes changed and the latter were deemed to be inferior goods. Thus, there are a number of ways in which tastes may change over time.

The next set of variables, the A variable, relates to levels of advertising, representing the level of own product advertising, the advertising of substitutes and the advertising of complements respectively. The relationships here are as follows. In general, the higher the level of own advertising for a good, the higher demand for that good would be expected, other things being constant. Likewise, the higher the level of advertising of a complementary good, the higher the demand for it and the good(s) which it is complementary to will be, given their symbiotic relationship. Conversely, however, the higher the level of advertising of a substitute good, the lower the demand for the good for which it is an alternative and people buy more heavily promoted good. The overall effect of advertising will depend on the extent to which each of these forms of advertising is used at any given point of time as they may, at least in part, cancel each other out. This is something the firm will also need to know in order to determine its optimal advertising strategy. The variables CR and R are also related. The former represents the availability of credit while the latter represents the rate of interest, that is the price of credit. These variables will be most important for purchases of consumer durable goods, for example cars. Someone's ability to buy a car will depend on his or her ability to raise money to pay for it. This means that the easier credit is to obtain, the more likely they are to be able to make the purchase. At the same time credit must be affordable, that is the rate of interest must be such that they have the money to pay. These two variables have traditionally been regarded as exogenous to the firm that is, they cannot be 'controlled' by it. In recent years, however, major car manufacturers have increasingly sought to bring them under their control through the provision of finance packages.

The letter E in the demand function stands for expectations. This may include expectations about price and income changes. For example, if consumers expect the price of a good to rise in future then they may well bring forward their purchases of it in order to avoid paying the higher price. This creates an increase in demand in the short term, but over the medium term, demand may fall in response to the higher price charged. The firm will need to adjust its production accordingly. An example of this might be when increased taxes are expected to be levied on particular goods, for example an increase in excise duties on alcohol or cigarettes, as is usually the case after the Central Budget. Consumers of these products may buy more of them prior to the implementation of the duty increases in order to avoid paying the higher prices arising from the higher level of duties. Alternatively, expectations about incomes may be important. For example, people who expect their incomes to rise may buy more goods, whereas those who expect their incomes to fall will buy less. At the level of the individual consumer this may not be significant but when aggregated across a country's population it can be. Thus during a boom in the economy the additional expected purchasing power of consumers will lead to increases in demand for a significant number of products. Conversely, the expectation that incomes will fall, perhaps as a

result of redundancy during a recession, will reduce demand as consumers become more cautious.

The variable N stands for the number of potential customers. Each product is likely to have a target market, the size of which will vary. The number of potential customers may be a function of age or location. For example, the number and type of toys sold in a particular country will be related to its demographic spread, in this case the number of children within it and their ages.

Finally, we come to O which represents any other miscellaneous factors which may influence the demand for a particular product. For example, it could be used to represent seasonal changes in demand for a particular product if demand is subject to such fluctuations rather than spread evenly throughout the year. Examples of such products might include things such as umbrellas, ice creams and holidays. In sum, this is a 'catch all' variable which can be used to represent anything else which the decision maker believes to have an effect on the demand for a particular product.

Thus each product will have its own particular demand function depending on which of the above variables influence the demand for it. The ways in which the level of demand can be estimated on the basis of this demand function will be discussed later in unit 6 of this Block.

Activity 1

Name any five variables which may be included in the demand function and explain its impact on demand.

.....

.....

.....

.....

4.3 THE LAW OF DEMAND

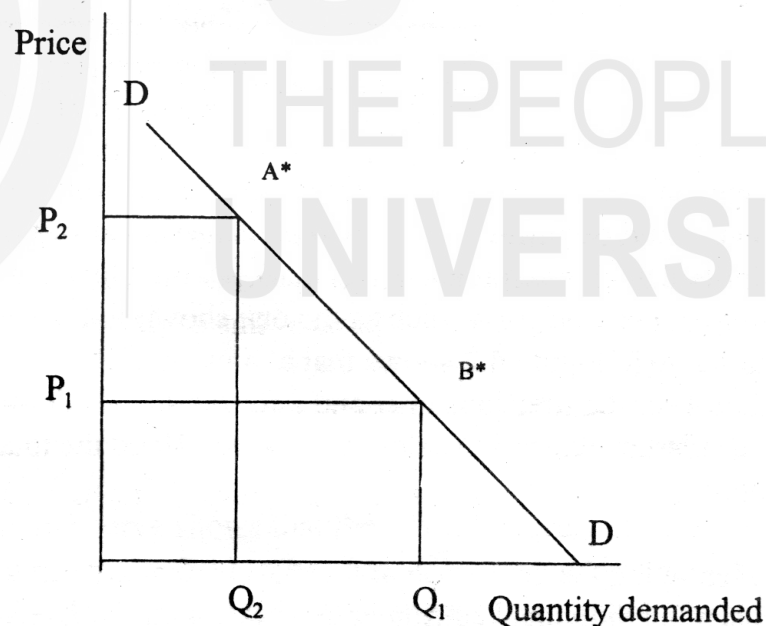
For most goods, consumers are willing to purchase more units at a lower price than at a higher price. The inverse relationship between price and the quantity consumers will buy is so widely observed that it is called the *law of demand*. The law of demand is the rule that people will buy more at lower prices than at higher prices if all other factors are constant. This idea of the law of demand seems to be a pretty logical and accurate description of the behaviour we would all expect to observe and for now, this will suffice.

The law of demand states that consumers are willing and able to purchase more units of a good or service at lower prices than at higher prices, other things being equal. Have you ever thought about why the law of demand is

true for nearly all goods and services? Two influences, known as the income effect and the substitution effect, are particularly important in explaining the negative slope of demand functions. The income effect is the influence of a change in a product's price on real income, or purchasing power. If the price of something that we buy goes down, our income will go up farther and we can purchase more goods and services (including the goods for which price has fallen) with a given level of money income. The *substitution effect* is the influence of a reduction in a product's price on quantity demanded such that consumers are likely to substitute that good for others that have thus become relatively more expensive.

The concept of demand is often depicted in a graphic model as a demand curve. A demand curve is a graphic illustration of the relationship between price and the quantity purchased at each price. When plotting a graph for demand, the price is measured along the vertical axis and the quantities that would be purchased at various prices are measured along the horizontal axis. The demand curve shows the relationship between the own price of a good and the quantity demanded of it. Any change in own price causes a movement along the curve as shown in Figure 4.1. In this case, a rise in price from P_1 to P_2 results in a fall in quantity demanded from Q_1 to Q_2 i.e. a move from B^* to A^* in the figure.

Figure 4.1: The demand curve



The same information can also be given in a table or demand schedule, such as Table 4.1, or by an equation for the demand function such as the following:

$$P = 100 - 0.25Q$$

where P is price and Q is quantity. The advantage of the equation is that it is compact to work with, and modern managers in both the private and public sector rely on such functions (which are estimated by using regression analysis, explained in unit 6) with increasing frequency.

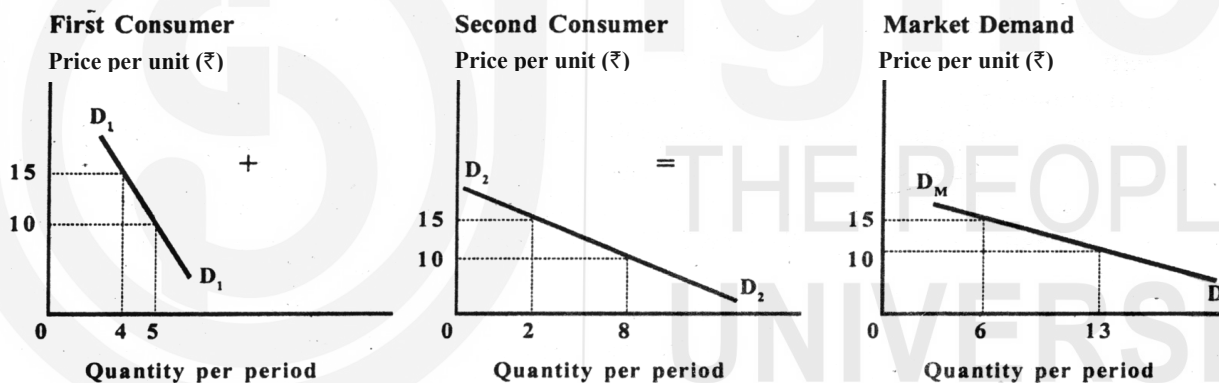
Table 4.1: Demand Schedule

Price (₹)	Quantity (units)
90	40
70	120
50	200
30	280
10	360

4.4 THE MARKET DEMAND CURVE

The market demand curve is the total of the quantities demanded by all individual consumers in an economy (or market area) at each price. Economic theory supports the proposition that individual consumers will purchase more of a good at lower prices than at higher prices. If this is true of individual consumers, then it is also true of all consumers combined. This relationship is demonstrated by the example in Figure 4.2, which shows two individual demand curves and the market demand that is estimated by adding the two curves together.

Figure 4.2 The Market Demand Curve



A market demand curve is the sum of the quantities that all consumers in a particular market would be willing and able to purchase at various prices. If we plotted the quantity that all consumers in this market would buy at each price, we might have a market demand curve such as the one shown in Figure 4.2. The market demand curve in Figure 4.2 shows that at a price of ₹ 15, the market demand would be 4 for the first consumer and 2 for the second consumer, giving a total of 6 units as market demand. Analogously, at ₹ 10.00 the total market demand is 13 units.

Another way of showing the derivation of the market demand curve is through equations representing individual consumer demand functions. Consider the following three equations representing three consumers' demand functions:

Consumer 1: $P = 12 - Q_1$

Consumer 2: $P = 10 - 2Q_2$

Consumer 3: $P = 10 - Q_3$

You should substitute some value of Q (such as $Q = 4$) in each of these equations to verify that they are consistent with the data in Table 4.2. Now, add these three demand functions together to get an equation for the market demand curve. Be careful while doing this. There is sometimes a temptation to just add equations without thinking about what is to be aggregated. In Table 4.2, it is easy to see that the quantities sold to each consumer at each price have been added. For example, at a price of ₹ 6, consumer number 1 would buy six units ($Q_1 = 6$), consumer number 2 would buy two units ($Q_2 = 2$), and consumer number 3 would buy four units ($Q_3 = 4$). Thus, the total market demand at a price of ₹ 6 is 12 units ($6 + 2 + 4 = 12$). The important point to remember is that the quantities are to be added; *not* the prices. To add the three given demand equations, we must first solve each for Q because we want to add the quantities (that is, we want to add the functions horizontally, so we must solve them for the variable represented on the horizontal axis). Solving the individual demand functions for Q as a function of P (for consumers 1, 2 and 3), we have—

$$Q_1 = 12 - P$$

$$Q_2 = 5 - 0.5P$$

$$Q_3 = 10 - P$$

Adding these equations results in the following:

$$Q_1 + Q_2 + Q_3 = 27 - 2.5P$$

And letting $Q_M = Q_1 + Q_2 + Q_3$ where Q_M is market demand.

$$Q_M = 27 - 2.5P$$

Q_M is the total quantity demanded.

This is the algebraic expression for the market demand curve. We could solve this expression for P to get the inverse demand function:

$$P = 10.8 - 0.4Q_M$$

Now, check to see that this form of expressing the market demand is consistent with the data shown in Table 4.2.

Table 4.2: Derivation of a Market Demand Schedule

Price	Q_1	Q_2	Q_3	Q_M
10	2	0	0	2
8	4	1	2	7
6	6	2	4	12
4	8	3	6	17
2	10	4	8	22

The market demand curve shows that the quantity purchased goes up from 12 to 22 as the price falls from ₹ 6.00 to ₹ 2.00. This is called *a change in*

quantity demanded. As the price falls, a greater quantity is demanded. As the price goes up, a smaller quantity is demanded. A *change in quantity demanded* is caused by a change in the price of the product for any given demand curve. This is true of individual consumers' demand as well as for the market demand. But what determines how much will be bought at each price? Why are more televisions bought now than ten years ago, despite higher prices? Why are more paperback books bought today than in previous years, even though the price has gone up? Questions such as these are answered by looking at the determinants of demand.

Activity 2

1. Given that an individual consumer's demand curve is $P = 200 - 4Q$.
 - a) Find the quantity this consumer would purchase at a price of Rs. 20.
 - b) Suppose that the price increases to ₹ 60. How much would the consumer now purchase?
 - c) Would this represent a change in demand or a change in quantity demanded? Why?

2. Suppose demand for a product in each of three regions is shown below:

Region 1 Demand: $Q = 307 - 5P$

Region 2 Demand: $Q = 204 - 3P$

Region 3 Demand: $Q = 500 - 9P$

Find the total demand for the three regions.

4.5 THE DETERMINANTS OF DEMAND

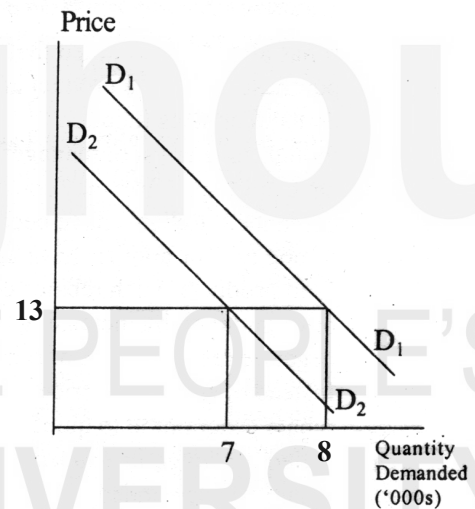
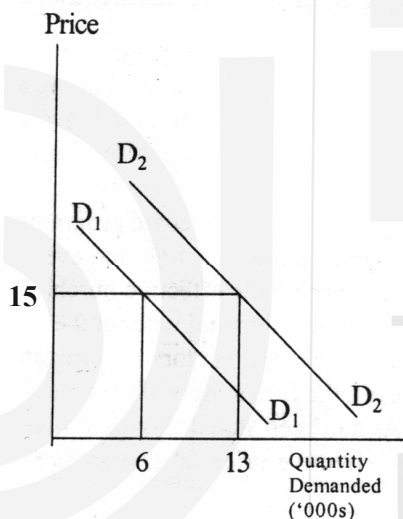
Many forces influence our decisions regarding the bundle of goods and services we choose to purchase. It is important for managers to understand these forces as fully as possible in order to make and implement decisions that enhance their firms' long-term health. It is probably impossible to know about all such forces, let alone be able to identify and measure them sufficiently to incorporate them into a manager's decision framework. However, a small subset of these forces is particularly important and nearly universally applicable. As stated above, the overall level of demand is

determined by consumers' incomes, their attitudes or feelings about products, the prices of related goods, their expectations, and by the number of consumers in the market. These are often referred to as the determinants of demand. **Determinants of demand** are the factors that determine how much will be purchased at each price. As these determinants change over time, the overall level of demand may change. More or less of a product may be purchased at any price because of changes in these factors.

Such changes are shown by a shift of the entire demand curve. If the demand curve shifts to the right, we say that there has been an *increase in demand*. This is shown as a move from the original demand D_1D_1 to the higher demand D_2D_2 in Figure 4.3a. The original demand curve can be thought of as being the market demand curve for soft drinks. At a price of ₹ 15 given the initial level of demand, consumers would purchase 6,000 soft drinks. If demand increases to the higher demand, consumers would purchase 13,000 soft drinks rather than the 6,000 along the original demand curve.

Figure 4.3a: Increased demand

Figure 4.3b: Decreased demand



A *decrease in demand* can be illustrated by a shift of the whole demand curve to the left. In Figure 4.3 b this is represented by a move from the original demand D_1D_1 to the lower demand D_2D_2 . At the price of ₹ 13 initially 8,000 soft drinks are purchased, while following the decrease in demand only 7,000 soft drinks are bought.

It is important to see that these changes in demand are different from the changes in quantity demanded. We discussed how changes in price cause a change in quantity demanded. As price changes, people buy more or less along a given demand curve. Movement from A^* to B^* in Figure 4.1 shows the change in quantity demanded as price changes. It is not a shift in the whole demand curve, such as that shown in Figure 4.3a and 4.3b. When the whole demand curve changes, there is a *change in demand*. Some of the variables that cause a change in demand are *changing incomes*, *changing tastes of consumers*, *changes in other prices*, *changes in consumer expectations*, and *changes in the number of consumers in the market* etc. These variables that cause a change in demand are also known as *shifter variables*.

Suppose you read in today's newspaper that carrot prices have soared because more carrots are being demanded. Then tomorrow you read that the rising price of carrots has greatly reduced the typical consumers demand for carrots as consumers have switched to potatoes and peas. The two statements appear to contradict each other. The first associates a rising price with rising demand; while the second associates a rising price with a declining demand. Comment.

.....

.....

.....

.....

4.6 THE PRODUCT'S PRICE AS A DETERMINANT OF DEMAND

It has already been noted that consumers are expected to be willing and able to purchase more of a product at lower prices than at higher prices. In evaluating a demand or sales function for a firm or an entire industry, one of the first things a thoughtful analyst or manager will consider is the price of a product. If inventories have built up, a firm may consider lowering the price to stimulate quantity demanded. Rebates have become a popular way of doing this. Rebate programmes of one type or another have appeared for cars, home appliances, toys and even food products. Such rebates constitute a way of lowering the effective purchase price and thereby increasing the quantity that consumers demand without the negative repercussions of raising the price once the excess inventory is eliminated. Instead of raising the price back to its normal level, the firm simply allows the rebate programme to quietly come to an end. As has been stated above, this is called a *change in quantity demanded*. As the effective price falls, a greater quantity is demanded.

4.7 INCOME AS A DETERMINANT OF DEMAND

On the other hand, shifter variables, as the name implies cause the demand curve to shift i.e., there is a change in demand. Nearly all goods and services are what economists refer to as *normal goods*. These are goods for which consumption goes up as the incomes of consumers rise, and the converse is also true. In fact, it is rare to find a demand function that does not include some measure of income as an important independent variable. Goods for which consumption increases as the incomes of consumers rise are called normal goods. Goods for which consumption decreases as the incomes of consumers rise are called *inferior goods*.

This relationship between product demand and income is one of the reasons that so much national attention is given to the level of Gross Domestic Product (GDP) and changes in the rate of growth of GDP. The GDP is the

broadest measure of income generated in the economy. In demand analysis, other more narrowly defined measures, such as personal income or disposable personal income, are often used; but these measures are highly correlated with GDP. Thus, looking at the changing trends in GDP is helpful for understanding what may happen to the demand for a product.

4.8 TASTES AND PREFERENCES AS DETERMINANTS OF DEMAND

We all like certain things and dislike others. A pair of identical twins brought up in the same environment may have different preferences in what they buy. Exactly how these preferences are formed and what influences them is not easy to know. Psychologists, sociologists, and social psychologists have a lot to offer in helping economists and other business analysts understand how preferences are formed and altered.

Even if we do not have a thorough understanding of preference structures, one thing is clear. Preferences and changes in preferences affect demand for goods and services. All have observed how such changes in tastes and preferences have influenced various markets. For example, consider the automobile market. In the United States, people appeared to have a preference for big, powerful cars throughout the 1950s and 1960s. During the 1970s, the preference structure started to change in favour of smaller, less-powerful, but more fuel-efficient cars. In part, the change in preference structure for cars may also have been related to lifestyle factors, such as being sportier and more concerned with resource conservation.

Convenience factors, such as ease of driving and parking, may also have been important. Demographic changes, especially a trend toward smaller families, may have had some effect as well. In terms of the theory, the change in preference toward fuel-efficient cars will shift the demand curve for smaller cars to the right (see Figure 4.3a). On the other hand, social attitudes towards smoking have changed and thus one would expect that the demand curve for cigarettes has shifted to the left (see Figure 4.3b). Likewise, the growing awareness in respect of noise and environmental pollution has resulted in a decline in the demand for crackers during Diwali celebrations.

As much as we may like to think that we know our own minds and make our own purchase decisions without the influence of others, we are very likely influenced quite strongly by various peer groups, including the people with whom we work, classmates, roommates, the people with whom we socialize, neighbours, and so on. Our decisions about clothes, entertainment, college courses, food, and many other things are influenced by these peer groups.

From the business perspective, advertising is a key factor in the formation and alteration of consumers' tastes and preferences. We can think of advertising as being either primarily informational or primarily transformational. Informational advertisements are designed to increase demand for a particular product by providing information about a product: how it is used, how much it costs, where you can buy it, what attributes it has (size, weight, etc.). Newspaper advertisements often have the strongest

informational component. They are designed to elicit direct purchase action on the part of consumers.

Transformational advertisements are designed to influence the image of a product or service. They attempt to enhance the satisfaction we get from a product by evoking positive images of how we will feel if we buy it. Think about television or social media commercials for products. How much usable information do they give you? Usually not much. They tell you the name of the product; associate its consumption with some reference group of happy, healthy-looking people (often famous people as well); and generally, try to create a positive image in an attempt to alter your preference structure in favour of that product. A single exposure to such an advertisement is not usually very effective, but repeated exposure to that advertisement and similar ones has been shown to influence consumer preferences.

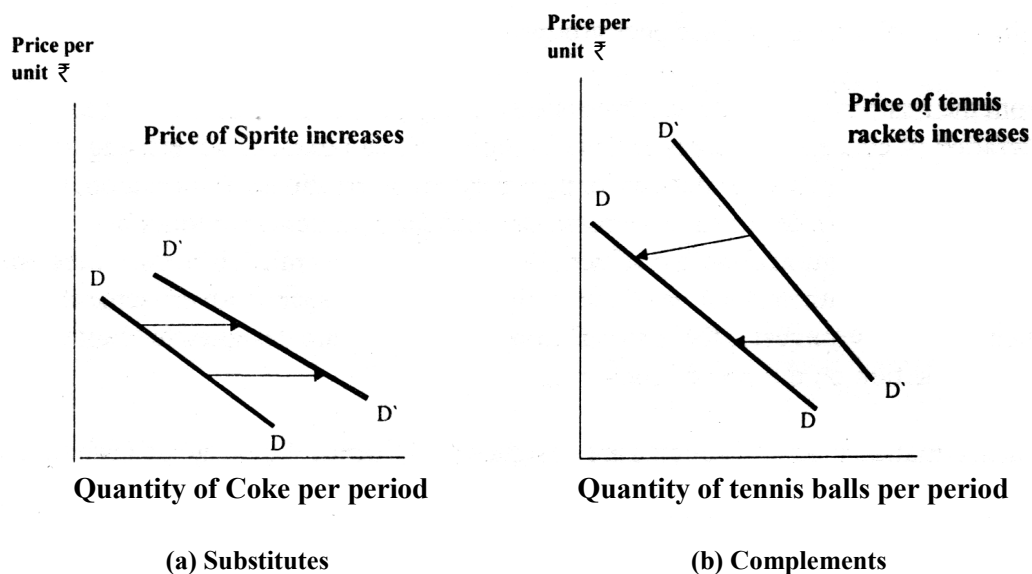
In terms of measurement, this determinant of demand-tastes and preferences is the most difficult for an analyst to handle. It is hard to identify all the things that influence tastes and preferences, and often just as hard to measure those that are identified. Measures of advertising effort are sometimes used to help account for this determinant of demand.

4.9 OTHER PRICES AS DETERMINANTS OF DEMAND

How much consumers buy of a product may be affected by the prices charged for other goods or services as well. Figures 4.4a and 4.4b show the effect on the demand curve following a change in the price of a related good or service. Both graphs are self-explanatory. Earlier, it was noted that the rise in the price of gasoline during the 1970s had some effect on the demand for large versus small cars in the United States. Gasoline and cars are complementary goods; they are used together and complement one another. When the price of gasoline rose, there were at least two effects on the automobile market. First, the higher price of gas increased the cost of driving, and thus reduced the total number of miles individuals tended to drive. Second, smaller, more fuel-efficient cars became more attractive relative to big cars.

This relationship can be stated in more general terms. Suppose that we observe two goods, A and B, and B is complementary to A. If the price of B goes up, we can expect the quantity demanded for A to be reduced. Why? Because as the price of goods B increases, its quantity demanded decreases according to the law of demand. But now, some individuals who would have purchased B at the lower price are no longer making those purchases. These same individuals now no longer have any use for A, because A was a good useful only in conjunction with B. Thus, the quantity demanded of A goes down as well. The reverse is also true: if the price of B falls, the demand for A will rise. It should be clear why business analysts are concerned not only about the effect that their product's price has on sales but also with the effect of the prices of complementary products.

Figure 4.4: Demand Curves for Substitutes and Complements



What effect would you expect of an increase in the price of tea to have on the demand for coffee? These are substitute goods. That is, we assume that people view tea and coffee in much the same manner; some people would be willing to drink either depending upon the price of the two drinks, the convenience of obtaining each, and so on. If the price of tea goes up, we would expect the demand for coffee to rise as well. Why? If the price of tea rises, fewer people drink tea (according to the law of demand). Some of them will now substitute coffee for the tea they would have drunk, and this is seen as an increase in the quantity demanded of coffee. In general, if we have two products, *C* and *D*, which are substitutes, we can expect that a rise in the price of *C* (or *D*) will cause the demand for *D* (or *C*) to go up.

Economists and other business analysts are, therefore, concerned with all other prices that may affect the products they are analyzing. The prices of both complementary and substitute products can be expected to influence demand. Later in this block, we will see how a measure called the cross-price elasticity of demand can be helpful in determining whether two goods are complements, substitutes or neither.

4.10 OTHER DETERMINANTS OF DEMAND

It would be a monumental task to identify everything that might have some influence on the demand for any product. So far, the four most important influences have been identified: a product's price, income, tastes and preferences, and the price of complementary or substitute products. A number of others were identified in section 4.2, which also affect demand. By now you will be able to establish the direction of the effect i.e., whether it will increase or decrease demand. For example, population growth obviously causes the potential demand for nearly all products to rise. In many cases, economists involved in analyzing the demand for particular products look at individual components of the population as determinants of demand. The changing age distribution, for example, may have differential effects on different markets. The growing proportion of people over 65 in the

population has important ramifications for demand for such things such as health-care products. Changes in other demographic characteristics and in the geographical distribution of the population may also be important. You may think of a variety of other effects on consumer demand as well.

4.11 SUMMARY

Demand analysis is a necessary informational input into the business decision process since, in a sense, demand fundamentally determines what is to be produced and at what price. Accordingly, business economists use demand analysis to discover the various factors determining the demand for a given product or service.

The law of demand states that consumers are willing and able to purchase more units of a good or service at lower prices than at higher prices, other things being constant.

The law of demand arises from two effects occurring when the price of a given good is changed: income and substitution. The *income effect* states that, given a fixed budget, increase in a product's price leaves less income for other goods. The *substitution effect* says that consumers tend to substitute cheaper goods for more expensive goods.

The market demand curve is the horizontal summation of all consumers demand in the market. The market demand curve relates the total quantity demanded of a product to its own price on the assumption that all other prices, total income of the individuals buying in that market, and its distribution, tastes and all other influencing variables are held constant.

The total quantity demanded in any market depends upon the price of the product being sold, on the prices of all other products, on the income of the individuals buying in that market, on the distribution of income among the individuals, tastes and a number of other influencing variables.

4.12 KEY WORDS

Law of demand states that the demand for a good varies inversely with its own price.

Market demand is the total of the quantities demanded by all individual consumers in an economy.

Substitutes are goods which can be used in place of the other goods.

Complements are goods used in conjunction with each other.

4.13 SELF-ASSESSMENT QUESTIONS

1. Give three reasons as to why the firm might need accurate demand information.

2. What causes a movement along the demand curve and what causes shifts in the demand curve? Explain.
3. Punita spends all her money on food and clothing. When the price of clothing decreases, she buys more clothing.
 - a. Does the substitution effect cause her to buy more clothing? Explain.
 - b. Does the income effect cause her to buy more clothing? Explain.
4. In a world of just two goods where all income is spent on the two goods, both goods cannot be inferior. True or False? Explain.

4.14 FURTHER READINGS

Allen, W.B., Weigelt, K. Doherty, N., & Mansfield, E. (2012), *Managerial Economics: Theory, Applications and Cases* (8th Ed.). W W Norton & Co.

Lewis, W.C., Jain, S.K., & Petersen, H.C. (2005). *Managerial Economics* (4th Ed.). Pearson.

UNIT 5 DEMAND ELASTICITY

Objectives

After studying this unit, you should be able to:

- **understand** the concept of elasticity;
- **define** various kinds of elasticities;
- **apply** the various measures of elasticity to real world situations.

Structure

- 5.1 Introduction
- 5.2 The Price Elasticity of Demand
- 5.3 Arc Price Elasticity
- 5.4 Point Price Elasticity
- 5.5 Price Elasticity and Revenue
- 5.6 Determinants of Price Elasticity
- 5.7 Income Elasticity of Demand
- 5.8 Cross-Price Elasticity
- 5.9 The Effect of Advertising on Demand
- 5.10 Summary
- 5.11 Key Words
- 5.12 Self-Assessment Questions
- 5.13 Further Readings

5.1 INTRODUCTION

In the previous unit, we studied that when price falls, quantity demanded would increase. While we know this qualitative effect exists for most goods and services, managers and business analysts are often more interested in knowing the magnitude of the response to a price change i.e. by how much? There are many situations in which one might want to measure how sensitive the quantity demanded is to changes in a product's price. Economists and other business analysts are frequently concerned with the responsiveness of one variable to changes in some other variable. It is useful to know, for example, what effect a given percentage change in price would have on sales. The most widely adopted measure of responsiveness is *elasticity*. Elasticity is a general concept that economists, business people, and government officials rely on for such measurement. For example, the finance minister might be interested in knowing whether decreasing tax rates would increase tax revenue. Likewise, it is often useful to measure the sensitivity of changes in demand to changes in one of the determinants of demand, such as income or advertising.

Elasticity is defined as the ratio of the percentage change in quantity demanded to the percentage change in some factor (such as price or income) that stimulates the change in quantity. The reason for using percentage change is that it obviates the need to know the units in which quantity and price are measured. For example, quantity could be in kilograms, grams, litres or gallons and price could be in dollars, rupees, euro etc. A measure of

elasticity based on units would lead to confusion and misleading comparisons across different products. The use of percentage change makes the measure of elasticity independent of units of measurement and hence easy to understand.

Elasticity is the percentage change in some dependent variable given a one-percent change in an independent variable, *ceteris paribus*. If we let Y represent the dependent variable, X the independent variable, and E the elasticity, then elasticity is represented as

$$E = \% \text{ change in } Y / \% \text{ change in } X$$

There are two forms of elasticity: arc elasticity and point elasticity. The former reflects the average responsiveness of the dependent variable to changes in the independent variable over some interval. The numeric value of arc elasticity can be found as follows:

$$\begin{aligned} \text{Arc elasticity} &= \frac{\text{change in } Y / \text{average } Y}{\text{change in } X / \text{average } X} = \frac{Y_2 - Y_1 / 0.5(Y_2 + Y_1)}{X_2 - X_1 / 0.5(X_2 + X_1)} \\ &= \frac{Y_2 - Y_1 / 0.5(Y_2 + Y_1)}{X_2 - X_1 / 0.5(X_2 + X_1)} \end{aligned}$$

where the subscripts refer to the two data points observed, or the extremes of the interval for which the elasticity is calculated.

Point elasticities indicate the responsiveness of the dependent variable to the independent variable at one particular point on the demand curve. Point elasticities are calculated as follows:

$$(e) = \frac{\partial Y}{\partial X} * \frac{X_1}{Y_1}$$

This form works well when the function is bivariate: $Y = f(X)$. However, when there are more independent variables, partial derivatives must be used. For example, suppose that $Y = f(W, X, Z)$ and we want to find the elasticities for each of the independent variables. We would have

$$e_w = \frac{\partial Y}{\partial W} * \frac{W}{Y}$$

$$e_x = \frac{\partial Y}{\partial X} * \frac{X}{Y}$$

$$e_z = \frac{\partial Y}{\partial Z} * \frac{Z}{Y}$$

Although economists use a great variety of elasticities, the following three deserve particular attention because of their wide application in the business world: price elasticity, income elasticity, and cross-price elasticity. We discuss these in detail in the subsequent sections.

5.2 THE PRICE ELASTICITY OF DEMAND

Price elasticity of demand measures the responsiveness of the quantity sold to changes in the product's price, *ceteris paribus*. It is the percentage change in sales divided by a percentage change in price. The notation E_p will be used for the arc price elasticity of demand, and e_p will be used for the point price elasticity of demand. If the absolute value of E_p (or e_p) is greater than one, a given percentage decrease (increase) in price will result in an even greater percentage increase (decrease) in sales¹. In such a case, the demand for the product is considered *elastic*; that is, sales are relatively responsive to price changes. Therefore, the percentage change in quantity demanded will be greater than the percentage change in the price. When the absolute value of the price elasticity of demand is less than one, the percentage change in sales is less than a given percentage change in price. Demand is then said to be *inelastic* with respect to price. *Unitary price elasticity* results when a given percentage change in price results in an equal percentage change in sales. The absolute value of the coefficient of price elasticity is equal to one in such cases. These relationships are summarized as follows:

If $|e_p|$ or $|E_p| > 1$, demand is elastic

If $|e_p|$ or $|E_p| < 1$, demand is inelastic

If $|e_p|$ or $|E_p| = 1$, demand is unitarily elastic

5.3 ARC PRICE ELASTICITY

Consider the hypothetical prices of some product and the corresponding quantity demanded, as given in Table 5.1. We could calculate the arc price elasticity between the two lowest prices i.e. between Rs. 30 and Rs. 10 as follows:

$$E_p = \frac{(360 - 280)}{(360 + 280)} \div \frac{(10 - 30)}{(10 + 30)} = -.25$$

Thus, demand is inelastic in this range. This value of $E_p = -.25$ means that a one percent change in price results in a .25% change in the quantity demanded (in the opposite direction of the price change) over this region of the demand function.

¹You should note that since the demand curve is downward sloping, i.e. price and quantity are inversely related; the price elasticity of demand will always be negative. Thus the change in quantity will be in the opposite direction to the change in price. We usually ignore the negative sign and consider absolute values for price elasticity to ease understanding of the concept.

Table 5.1: Demand Schedule to Demonstrate Price Elasticities

Price ₹ (P)	Quantity (units) (Q)	Arc Elasticity	Point Elasticity
90	40	– 4.00	– 9.00
70	120	– 1.50	– 2.33
50	200	-0.67	-1.00
30	280	– 0.25	– 0.43
10	360		– 0.11

If we calculate the arc price elasticity between the prices of 50 and 70, we have

$$E_p = \frac{(200 - 120)}{(200 + 120)} / \frac{(50 - 70)}{(50 + 70)} = -1.5$$

We would say that demand is price elastic in this range because the percentage change in sales is greater than the percentage change in price. You can calculate arc elasticity over any price range. As an exercise estimate the arc elasticity between the extremes of the demand function shown in Table 5.1, i.e. between ₹ 90 and ₹ 10. Satisfy yourself that the absolute value of arc elasticity between these two points is 1.

5.4 POINT PRICE ELASTICITY

The algebraic equation for the demand schedule given in Table 5.1 is

$$P = 100 - .25Q$$

$$\text{or } Q = 400 - 4P$$

We can use this demand function to illustrate the determination of point price elasticities. Let's select the point at which $P = 10$ and $Q = 360$:

$$e_p = \frac{dQ}{dP} * \frac{P}{Q}$$

$$e_p = (-4)(10/360)$$

$$e_p = -.11$$

Because $|e_p| < 1$, we would say that demand is inelastic at a price of ₹ 10. Now, consider a price of ₹ 70:

$$e_p = \frac{dQ}{dP} * \frac{P}{Q}$$

$$e_p = (-4)(70/120)$$

$$e_p = -2.33$$

Here, $|e_p| > 1$, and demand is price elastic.

This example shows that the price elasticity of demand may (and usually does) vary along any demand function, depending on the portion of the function for which the elasticity is calculated. It follows that we usually cannot make such statements as "the demand for product X is elastic" because it is likely to be elastic for one range of prices and inelastic for another. Usually at high prices demand is elastic, while at lower prices demand tends to be inelastic. Intuitively, this is so because lowering price from very high levels is likely to stimulate demand much more compared to lowering prices when price is already low. As an illustration, consider the prices of cellular phones (handsets) when these were first introduced in the Indian market at prices ranging between ₹ 25,000 to ₹ 30,000 per handset. Demand was limited to the higher end of the market. As these prices fell, demand was stimulated and resulted in increasing penetration in the middle and lower end segments, indicating an elastic response.

Activity 1

1. What is the generic equation used to calculate a point elasticity of demand?

.....

.....

.....

.....

2. Use this equation to calculate the following elasticities:

- (a) The own price elasticity of demand for cars when the original price of a car of ₹ 8 lakhs is raised to ₹9 lakhs, resulting in a fall in the number of cars sold from 50,000 to 40,000.

.....

.....

.....

.....

- (b) The own advertising elasticity of demand when a firm increases its spending on advertising from ₹10,000 to ₹15,000, which leads to a rise in the number of chocolate bars sold from 50,000 to 75,000.

.....

.....

.....

.....

5.5 PRICE ELASTICITY AND REVENUE

We have defined demand to be elastic when the absolute value of the price elasticity is greater than one. For that to be true, the percentage change in quantity must be greater than the percentage change in price (% change in $Q > \% \text{ change in } P$). If this were true, what would you expect to happen to a firm's receipts if the price were lowered? Recall from principles of economics that total revenue (TR) is equal to price (P) times quantity (Q).

Consider an extreme case. Suppose that a five-percent cut in price stimulates a fifty-per cent increase in sales (the price elasticity would be 10). You would expect revenues to rise. The relatively small drop in price would be more than compensated for by a large increase in sales. To see exactly the relationship between total revenue and price elasticity, let's return to the demand function given by the equation $Q = 400 - 4P$ which is used as the basis for Table 5.2. The table shows how the price elasticity of demand varies along the demand curve.

Table 5.2: Demand, Price Elasticity and Revenue

Price (P)	Quantity (Q)	Total Revenue ($TR = P \cdot Q$)	Marginal Revenue ($MR = dTR/dQ$)	Point Elasticity $e_p = dQ/dP \cdot P/Q$
100	0	0	100	
90	40	3,600	80	- 9.00
80	80	6,400	60	- 4.00
70	120	8,400	40	- 2.33
60	160	9,600	20	- 1.50
50	200	10,000	0	- 1.00
40	240	9,600	- 20	- 0.67
30	280	8,400	- 40	- 0.43
20	320	6,400	- 60	- 0.25
10	360	3,600	- 80	- 0.11

Note that no point elasticity can be calculated when $Q = 0$ because division by zero is not defined.

In this table, total revenue and marginal revenue are included, as well as the point price elasticities. Marginal revenue (MR) is defined as the rate of change in total revenue, or the additional revenue generated by selling one more unit. In this example, the demand function can be solved for P in terms of Q as follows:

$$\begin{aligned} \text{Consider } Q &= 400 - 4P \\ \Rightarrow 4P &= 400 - Q \\ \Rightarrow P &= 100 - 0.25Q \end{aligned}$$

Multiplying by Q and taking the first derivative yields:

$$\begin{aligned} TR &= P \cdot Q \\ TR &= (100 - 0.25Q)Q \end{aligned}$$

$$TR = 100Q - 0.25Q^2$$

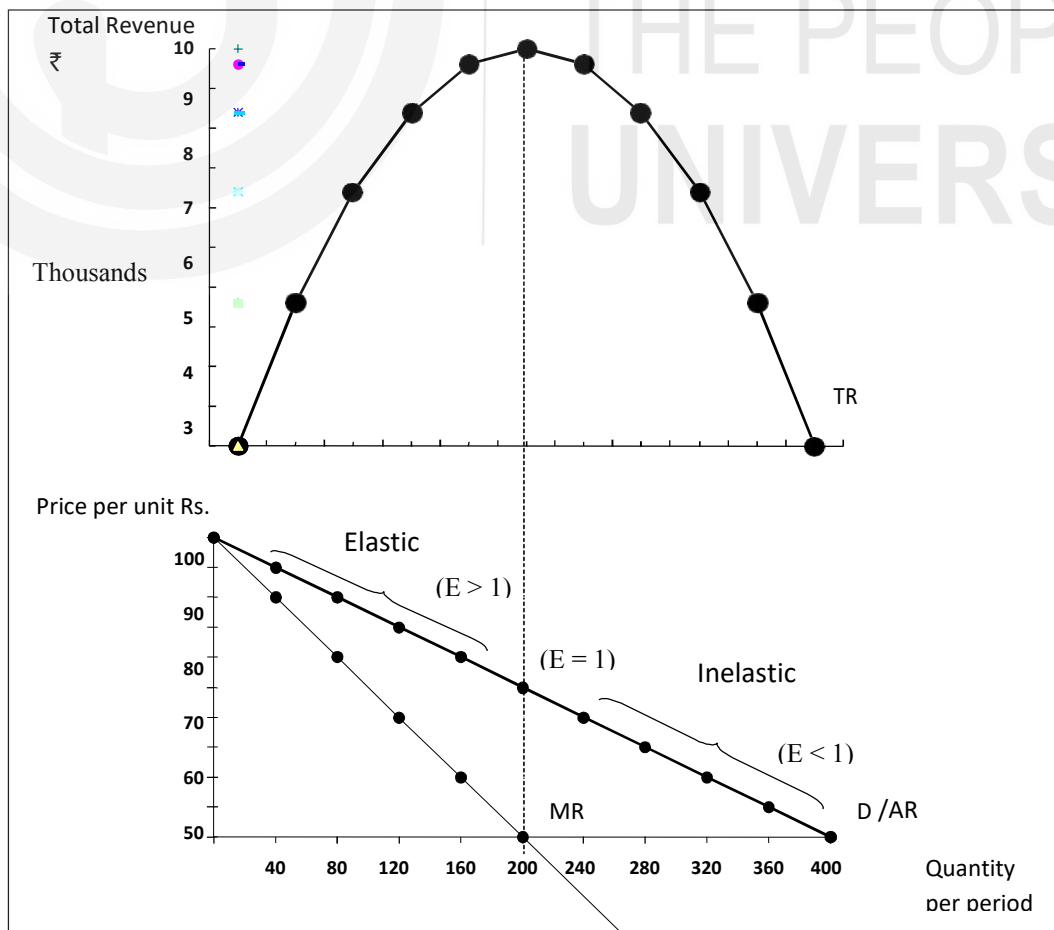
$$MR = dTR/dQ$$

$$MR = 100 - 0.5Q$$

The total revenue and marginal revenue functions along with the demand curve, are plotted in Figure 5.1. Notice that the slope of the marginal revenue function is twice the slope of the demand function. You see in Figure 5.1 that demand is price elastic over the range of quantities for which marginal revenue is positive. Because marginal revenue is the slope of total revenue (remember that $MR = dTR/dQ$), you can tell that increasing sales by lowering price will cause total revenue to rise over this interval. However, lowering price when demand is inelastic (beyond $Q = 200$) will result in reduced revenues. From an examination of Figure 5.1 and Table 5.2, we can reach some important conclusions about the relationship between elasticity and total revenue.

As is evident from the above discussion, the change in expenditure when price changes is related to the elasticity of demand. If elasticity is less than unity (inelastic), the percentage change in price can exceed the percentage change in quantity.

Figure 5.1: Illustrates the relationship between a firm's total revenue and the firm's demand curve. Over the range between 0 and 200 units, the demand function is elastic; over this same range, total revenue increases as price is reduced and quantity demanded increases. Beyond 200 units, demand is inelastic and total revenue decreases as price is reduced and quantity demanded increases.



The price change will then be the dominant one of the two changes and the revenue will change in the same direction as the price change. If however, elasticity exceeds unity (elastic), the percentage change in quantity will exceed the percentage change in price. The percentage change in quantity will be the more important change, so that total expenditure will change in the opposite direction as the price change. These results can be summarized as follows:

Elastic Demand

1. Decrease price..... Increase total revenue
2. Increase price.....Decrease total revenue

Price and total revenue move in opposite directions.

Inelastic Demand

1. Decrease price..... Decrease total revenue
2. Increase price.....Increase total revenue

Price and total revenue move in the same direction.

If a demand function has a unitary elasticity, then the same level of revenue will be generated, regardless of price. You see that for a linear demand function, as price falls, demand becomes less elastic or more inelastic. You have also seen that when demand is elastic, price cuts are associated with increases in total revenue. But if price continues to be lowered in the range in which demand is inelastic, total revenue will fall. Thus, total revenue will be maximized at the price (and related quantity) at which demand is unitarily elastic. You can see from Figure 5.1 that this point is also where $MR = 0$. This observation makes sense because marginal revenue is the slope of total revenue. Recall that any function may have a maximum where its slope is zero. For total revenue, we have

$$TR = P \cdot Q = (100 - 0.25Q) Q$$

$$TR = 100Q - 0.25Q^2$$

$$dTR/dQ = 100 - 0.5Q$$

Setting the first derivative equal to 0 and solving for Q, we find

$$dTR/dQ = 100 - .5Q = 0$$

$$-.5Q = -100$$

$$Q=200$$

Checking the second-order condition, we see that

$$d^2TR/dQ^2 = -0.5 < 0$$

Because the second derivative is negative, we know that total revenue is a maximum at $Q = 200$. The price at which 200 units will sell is Rs. 50, so that is the revenue-maximizing price.

The relationship between elasticity and total revenue can also be shown using calculus. Total revenue is price times quantity. Taking the derivative of total revenue with respect to quantity yields marginal revenue:

$$TR = P \cdot Q$$

$$MR = \frac{d(TR)}{dQ} = \frac{d(PQ)}{dQ} = P + Q \frac{dP}{dQ}$$

The equation states that the additional revenue resulting from the sale of one more unit of a good or service is equal to the selling price of the last unit (P), adjusted for the reduced revenue from all other units sold at a lower price (QdP/dQ). This equation can be written

$$MR = P \left(1 + \frac{QdP}{PdQ} \right)$$

But note that $(Q/P) dP/dQ = 1/e_p$

$$MR = P \left(1 + \frac{1}{e_p} \right)$$

This equation indicates that marginal revenue is a function of the elasticity of demand. For example, if demand is unitary elastic, $e_p = -1$ then,

$$MR = P \left(1 + \frac{1}{-1} \right) = 0$$

Because marginal revenue is zero, a price change would have no effect on total revenue. In contrast, if demand is elastic, say $e_p = -2$, marginal revenue will be greater than zero. This implies that a price reduction, by stimulating a considerable increase in demand would increase total revenue. This equation also implies that if demand is inelastic, say $e_p = -0.5$, marginal revenue is negative, indicating that a price reduction would decrease total revenue.

Some analysts question the usefulness of elasticity estimates. They argue that elasticities are redundant, such that the data necessary for their determination could be used to determine total revenues directly. Thus managers could assess the effects of a change in price without knowledge of price elasticity. Although this is true, elasticity estimates are valuable to the extent that they provide a quick way of evaluating pricing policies. For example, if demand is known to be elastic, it is also known that a price increase will reduce total revenues. Likewise, if the Finance Minister wants to raise revenue through taxation, he knows that increasing duties on cigarettes will most likely generate the result he desires.

The next section illustrates the usefulness of knowledge of elasticity.

Illustration

Price Elasticity of Demand: A Case Study of a Cigarette Company

A Swiss-American multinational cigarette and tobacco manufacturing company, with products sold in over 180 countries. In 2018, it cut its cigarette prices by 18%. Due to this decrease in price, their quantity sold increased by 12.5%, as there is an inverse relationship between price of the product and demand of the product. Later in 2019, it was found out that profits of Cigarette Company fell by 25% due to this bad pricing strategy. Now the question arises why profits fell despite increase in units sold. Here comes the concept of price elasticity of demand in picture.

If we calculate price elasticity of demand in this case, it is as follows:

$$e_p = \frac{\%Q}{\%P}$$

$$e_p = \frac{\%12.5}{\%18}$$

$$e_p = .69$$

It is less than one which means elasticity of demand is inelastic. Demand is inelastic, when % change in quantity sold is less than % change in price. In this case, demand for cigarettes is inelastic and with decrease in price, total revenue of the firm decreases as increase in units sold cannot offset the lower price per unit sold. So, total revenue decreases and total cost of production increases because more cigarettes were produced, which led to fall in profits. If Cigarette Company managers had paid attention to the economic concept of elasticity, they would have strategized better and profits would not have fallen.

Activity 2

1. If the marginal revenue is zero, then
 - a. Total revenue is zero
 - b. Average revenue is zero
 - c. Total revenue is at a maximum or a minimum
 - d. Average revenue is at a maximum
2. The total revenue function of an economics text book is given by $TR = 120Q - 0.1Q^2$.
 - a. Over what output range is demand elastic.
 - b. Initially price is set at Rs. 71.60. To maximize revenue should price be increased or decreased. Explain.

3. Complete the following Table:

Price	Sales	Total Revenue	Marginal	Arc Price Elasticity
3.50	80			
3.40	85			
3.30	90			
3.20	95			
3.10	100			
3.00	105			
2.90	110			
2.80	115			
2.70	120			
2.60	125			

5.6 DETERMINANTS OF PRICE ELASTICITY

Price elasticities can be estimated for many goods and services; Table 5.3 provides some examples. The short-run elasticities reflect periods of time that are not long enough for the consumer to adjust completely to changes in prices. The long-run values refer to situations where consumers have had more time to adjust. Note the variation in elasticities in Table 5.3. The long-run demand for foreign travel by U.S. residents is elastic (i.e., $e_p = |4.10|$). In contrast, the long-run demand for water is highly inelastic (i.e., $e_p = |0.14|$). Demand for is inelastic in the short run, but elastic in the long run. In general, three factors determine the price elasticity of demand. They are: (1) availability of substitutes, (2) proportion of income spent on good or service, and (3) length of time.

Availability of Substitutes: The main determinant of elasticity is the availability of substitutes. Products for which there are good substitutes tend to have higher price elasticity of demand than products for which there are few adequate substitutes.

Movies are a good example. Movies are a form of recreation, but there are many alternative recreational activities. When ticket prices at the movie theatre increase, these substitute activities replace movies. Thus, the demand for motion pictures is relatively elastic, as shown in Table 5.3. Other examples of products with close substitutes and therefore elastic demand would be demand for cars, demand for smart phones, demand for air-travel etc.

Table 5.3: Estimates of Price Elasticity

Good or Service	Estimated Price Elasticity	
Electricity	– 0.13	Short run
Electricity	– 1.89	Long run
Water	– 0.14	Long run
Motion pictures	– 3.69	Long run
Gasoline	– 0.15	Short run
Gasoline	– 0.78	Long run
Foreign travel	– 4.10	Long run

Sources: H. S. Houthakker and L. D. Taylor, *Consumer Demand in the United States. Analysis and Projections* (Cambridge, Mass.: Harvard University Press, 1970), and I. L. Sweeney, "The Demand for Gasoline: A Vintage Capital Model," Department of Engineering Economics, Stanford University, 1975.

At the other extreme, consider the short-run demand for electricity. When your local supplier increases prices, consumers have few options. There are not many short-run alternatives to using electricity for cooling and lighting. Hence the short-run demand for electricity is relatively inelastic. In the days of the license raj in India, when government was the monopoly provider, demand for telecom services was relatively inelastic since there was no other service provider in the market. Thus, a product with close substitutes tends to have elastic demand; one with no close substitutes tends to have inelastic demand. An important mission for most advertising is to make the consumer perceive that no close substitute exists for the product being advertised, thereby rendering the consumers demand relatively inelastic.

Proportion of Income Spent: Demand tends to be inelastic for goods and services that account for only a small proportion of total expenditures. Consider the demand for salt. 1kg of salt will meet the need of the typical household for a month and costs only a few rupees. If the price of salt were to double, this change would not have a significant impact on the family's purchasing power. As a result, price changes have little effect on the household demand for salt. In contrast, demand will tend to be more elastic for goods and services that require a substantial portion of total expenditures. Thus, demand for holiday travel and luxury cars take up a considerable portion of the family's budget and therefore tend to have higher elasticities. The relative necessity of a good also influences elasticity. For example, the demand for insulin is probably very inelastic because it is necessary for diabetics who rely on this drug.

Time Period: Demand is usually more elastic in the long run than in the short run. The explanation is that, given more time, the consumer has more opportunities to adjust to changes in prices. Table 5.3 indicates that the long-run elasticity for electricity is more than ten times the short-run value.

Price Elasticity and Decision Making

Information about price elasticities can be extremely useful to managers as they contemplate pricing decisions, if demand is inelastic at the current price, a price decrease will result in a decrease in total revenue. Alternatively, reducing the price of a product with elastic demand would cause revenue to increase. The effect on total revenue would be the reverse for a price increase. However, if demand is unitary elastic, price changes will not change total revenues. However, a price reduction is not always the correct strategy when demand is elastic. The decision must also take into account the impact on the firm's costs and profits. More will be said about pricing strategy in Block 4.

As another example of how knowledge of price elasticity may be useful, let's consider the demand for cigarettes. The price elasticity for cigarettes by age groups has been found to be:

Age Group	Price Elasticity
12-17 years	- 1.40
20-25 years	- 0.89
26-35 years	- 0.47
36-74 years	- 0.45

These elasticities indicate that young smokers are much more responsive to price than are older smokers. This may be in part related to the fraction of income that goes towards the purchase of cigarettes. It may also reflect the degree to which physical/psychological addiction influences consumption. From the perspective of cigarette sellers, these results suggest that if all sellers increased prices proportionately, the total expenditure on cigarettes by adult smokers would increase. (Recall that when demand is inelastic, price and total revenue move in the same direction). Individual brands would be more price elastic than for the entire product class because each brand has other brands that represent potential substitutes; however, for the product class, there may be few good substitutes.

Activity 3

- Several years ago a mayor of a large city was faced with a rapidly rising rate of theft, mugging and armed robbery by persons addicted to drugs. Since those who needed drugs were unable to earn enough money to purchase sufficient amount of drugs they were turning to crime. The mayor said that he strongly intended to reduce drug-related crime-mugging theft etc. The policy he adopted to solve this problem was to crack down strongly on the drug traffic. He arrested the drug dealers and reduced the amount of drugs on the street; but the rate of crime increased instead of decreasing. What economic explanation can you offer for the failure of the mayor's policy?

.....

.....

.....

.....

- The Automobile workers union once contended that automobile manufacturers should raise wages and simultaneously cut automobile prices. Arguing that the elasticity of demand for automobiles was about 4, the Union concluded that a price cut would help check inflation, boost the total receipts of producers and preserve or even increase the profits of the manufacturers. A spokesman of a automobile company however claimed that the elasticity of demand for automobiles was in the 0.5-1.5 range. In this case the elasticity of demand for automobiles become a strategic factor in labour-management relations and wage bargaining. Discuss.

5.7 INCOME ELASTICITY OF DEMAND

The income elasticity of demand measures the responsiveness of sales to changes in income, *ceteris paribus*. It is defined as the percentage change in sales divided by the corresponding percentage change in income. The methods used to calculate arc income elasticity (E_1) and point income elasticity (e_1) are as follows:

$$E_1 = \frac{Q_2 - Q_1}{I_2 - I_1} * \frac{I_2 + I_1}{Q_2 + Q_1}$$

$$e_1 = \frac{\delta Q}{\delta I} * \frac{I_1}{Q_1}$$

Given information on sales and income, the calculation of income elasticities is strictly analogous to the calculation of price elasticities. If the income elasticity of demand for a product is greater than one, the product is said to be income elastic; if it is less than one, the product is income inelastic. For normal goods, the income elasticity is greater than 0 because with rising incomes, consumers will purchase a greater quantity of such goods, *ceteris paribus*. If the income elasticity for a commodity is negative, the good is an inferior good; that is, people will choose to purchase less of the product when their income increases. Potatoes may represent examples of inferior goods for some households, as would purchases from the cheap stores. The reason is that some households consume certain goods only because of lack of purchasing power. As income increases it is possible the household will shift away from the purchase of these inferior goods. Income elasticity relationships can be summarized as follows:

Normal goods are indicated by e_1 or $E_1 > 0$. Inferior goods are indicated by e_1 or $E_1 < 0$.

If e_1 or $E_1 > 1$, the good is income elastic.

If e_1 or $E_1 < 1$, the good is income inelastic.

If e_1 or $E_1 = 1$, the good is unitarily income elastic.

To illustrate just one way in which income elasticity may be useful, consider the following situation. A firm has obtained a fairly reliable estimate of the projected percentage increase in income for its market area for the next year; let's say 4.5 per cent. Managers know that sales are currently running at an annual rate of 200,000 units, and the marketing analysis group has estimated the arc income elasticity of demand for the product at 1.2. If other factors are

expected to remain relatively constant, we can use this information as one input into projecting sales for the next year, as follows:

$$E_1 = \frac{\Delta\%Q}{\Delta\%I}$$

and subsequently

$$\% \Delta Q = (E_1)(\% \Delta I)$$

$$\% \Delta Q = (1.2)(4.5)$$

$$\% \Delta Q = 5.4$$

Thus, next year's sales would be projected to be 5.4 per cent above the current level, or 1.054 times this year's sales:

$$(1.054)(200,000) = 210,800 \text{ units}$$

Knowledge of income elasticities is also useful at different stages of a business cycle. For example, during periods of expansion, firms selling luxury products such as exotic vacations or big cars find demand for their products will increase at a rate that is faster than the rate of growth of incomes. However, during an economic recession demand may decrease rapidly for such products. Conversely, sellers of necessities such as basic food will not profit much during periods of economic prosperity, but will also find that their products are recession proof.

5.8 CROSS-PRICE ELASTICITY

The sales volume of one product may be influenced by the price of either substitute or complementary products. Cross-price elasticity of demand provides a means to quantify that type of influence. It is defined as the ratio of the percentage change in sales of one product to the percentage change in price of another product. The relevant arc (E_c) and point (e_c) cross-price elasticities are determined as follows:

$$E_c = \frac{Q_{b2} - Q_{b1}}{P_{a2} - P_{a1}} + \frac{P_{a2} + P_{a1}}{Q_{b2} + Q_{b1}}$$

$$e_c = \frac{\delta Q_a}{\delta P_b} * \frac{P_b}{Q_a}$$

where the alphabetic subscripts differentiate between the two products involved. A negative coefficient of cross-price elasticity implies that a decrease in the price of product A results in an increase in sales of product B, or vice versa, we can conclude that the products are complementary to one another (such as cassette tape players and cassette tapes). Thus, when the coefficient of cross-price elasticity for two products is negative, the products are classified as complements.

A similar line of reasoning leads to the conclusion that if the cross-price elasticity is positive, the products are substitutes. For example, an increase in the price of sugar would cause less sugar to be purchased, but would increase

the sale of sugar substitutes. When we calculate the cross-price elasticity for this case, both the numerator and the denominator (% change in Q of sugar substitutes and % change in P of sugar, respectively) would have the same sign, and the coefficient would be positive.

If two goods are unrelated, a change in the price of one will not affect the sales of the other. The numerator of the cross-price elasticity ratio would be 0, and thus the coefficient of cross-price elasticity would be 0. In this case, the two commodities would be defined as independent. For example, consider the expected effect that a 10% increase in the price of eggs would have on the quantity of electronic calculator sales.

These relationships can be summarized as follows:

If e_c or $E_c > 0$, goods are substitutes

If e_c or $E_c < 0$, goods are complementary

If e_c or $E_c = 0$, goods are independent

Cross price elasticities may not always be symmetrical. For example, consider two dailies, LOI and the HOI competing in the Delhi market. Most analysts will agree that the two products are substitutes i.e. the cross price elasticity is positive. However, there is no reason to believe that the change in demand for the LOI following a one percent change in the price of HOI will be equal to the change in demand for HOI following a one per cent change in the price of the LOI.

Many large corporations produce several related products. Automobile Company produces many varieties of automobiles, Detergent Company produces many brands of soap and Razor Company produces many types of razors. If Automobile Company reduces the price of its A-model, sales of its related S-model may decline. When a company sells related products, knowledge of cross elasticities can aid decision makers in assessing such impacts.

5.9 THE EFFECT OF ADVERTISING ON DEMAND

Advertising influences our attitudes towards the product or service being promoted. In most cases, the intent of a firm's advertising is to stimulate sales of a particular product or product line. When Pepsi Cola Corporation decides to sponsor a television show or cricket match it hopes that doing so will increase the sales of its products. Such product promotions have their impact on consumers through tastes and preferences.

In addition to shifting the demand function to the right, advertising may have the effect of making it somewhat more steep. The reason for this is that advertisements can create stronger consumer brand preferences, thus making consumers less sensitive to price changes for that product. This means that one effect of advertising can be to make the demand for a firm's product more price-inelastic. To the extent that this is true, management has an

increased ability to raise price without losing as many sales as would have been lost otherwise. We have seen that raising the product's price will increase total revenue for the firm if demand is inelastic.

5.10 SUMMARY

The concept of elasticity is simply the slope relationship of two variables expressed in percentage terms. This makes elasticity independent of the units and magnitudes of the two variables in question.

Price elasticity is an important determinant of the price firms will charge for their product. This is because price elasticity is related to the firm's marginal revenue function determining the effect on total revenue of a change in price.

When demand is price elastic, lowering price will increase total revenue; and when demand is inelastic, lowering price will decrease total revenue. The reverse is true for increase in price. When demand is price elastic, increasing price will decrease total revenue; and when demand is inelastic, increasing price will increase total revenue.

The main determinant of price elasticity of demand is the availability of substitutes for the product. Any one of a group of close substitutes will have a more elastic demand than the group as a whole. Elasticity will be low if the proportion of income spent on the product is small.

Elasticity tends to be greater, the longer the time over which adjustment occurs. Items that have few substitutes in the short run may develop substitutes when the consumers and producers have time to adapt.

Income elasticity of demand is the percentage change in quantity demanded as a result of a percentage change in income. Income elasticity will usually change as income varies.

Cross elasticity is the percentage change in quantity demanded divided by the percentage change in the price of some other product that brought it about. It is used to define products that are substitutes for one another (positive cross elasticity) and products that complement one another (negative elasticity).

5.11 KEY WORDS

Elasticity measures the responsiveness of one variable to the variations in another variable.

Price elasticity of demand measures the degree of responsiveness of the quantity demanded of a particular commodity to a change in price of that commodity.

Income elasticity of demand measures the responsiveness of demand for a commodity to a change in consumer's income.

Cross elasticity of demand measures the responsiveness of demand for one good to a change in price of the other good.

5.12 SELF-ASSESSMENT QUESTIONS

1. In early 1995, there was a sharp increase in the price of newsprint, the paper used by newspapers. Since newsprint is the second-largest expense for Indian newspapers (after salaries), publishers were concerned about the price hike. Suppose that the demand for newsprint can be represented as followings:

$$Q_1 = 17.3 - .0092P + 0.0067I,$$

where Q_1 equals the quantity demanded (in kilograms per capita), P is the price of newsprint (in ₹ per metric ton), and I is income per capita (in ₹).

(a) If there are 1 million people in the market, and if per capita income equals ₹10,000, what is the demand curve for newsprint? (b) Under these circumstances, what is the price elasticity of demand if the price of newsprint equals ₹400 per metric ton? (c) According to a 1994 study, demand curve for newsprint in the north eastern Indian state is

$$Q_2 = 2672 - 0.51P,$$

2. Where Q_2 is the number of metric tons of newsprint demanded (in thousands). What is the price elasticity of demand for newsprint in the north eastern Indian states if price equals ₹500 per metric ton? (d) Based on this study, will the 1995 price increase result in an increase or decrease in the amount spent on newsprint in the northeastern Indian States? Why?
3. The following table shows how sales (Q), total revenue (TR) and marginal revenue (MR) vary as prices change.

Price	Quantity	Total Revenue	Marginal Revenue
180	44	7920	—
160	48	7680	— 60
140	52	7280	— 100
120	56	6720	— 140
100	60	6000	— 180
80	64	5120	— 220
60	68	4080	— 260
40	72	2880	— 300
20	76	1520	— 340

- a. Calculate Arc price elasticities for the following price changes:
From $P = 180$ to $P = 160$:
From $P = 120$ to $P = 100$: From $P = 60$ to $P = 40$:
What can you say about price elasticity?
- b. Calculate point elasticity at $P = 100$

4. Consider the price and sales data in Activity 2, reproduced for easy reference.

Price	Sales
3.50	80
3.40	85
3.30	90
3.20	95
3.10	100
3.00	105
2.90	110
2.80	115
2.70	120
2.60	125

- Does this relationship follow the law of demand? Explain.
- Calculate the value of the arc price elasticity for a change in price from ₹3.40 to ₹3.30. Is demand elastic or inelastic in this region?
- Plot these 10 observations with price on the Y-axis and quantity on the x-axis connects them to form a demand curve.
- Based on this graph, how many units would you expect the firm to sell at a price of ₹2.95?
- The algebraic equation for the firms demand function based on the data is $P = 5.10 - .02Q$
- If the firm has a target sales level of 118 units, what price should be used?
- Find the algebraic expression for the firms' total revenue (TR) as a function of sales (Q).
- What is the algebraic equation for marginal revenue (MR)?
- Based on these algebraic functions, find the price and the corresponding level of sales that would maximize total revenue for the firm.
- Rewrite the firms demand function with Q as a function of P rather than P as a function of Q, as given earlier.
- What is the point price elasticity of demand at a price of ₹3.15? Is demand elastic or inelastic at this point?

5.13 FURTHER READINGS

Perloff, J.M.(2014). “*Microeconomics*”, Second Edition, Pearson Education Asia.

Chrystal, K.A., & Lipsey, R.G. (1997). *Economics for Business and Management*. Oxford University Press.

Alper, N.O., Archibald, R.B., and Jensen, E. (1987) “At What Price Vanity? An Econometric Model of the Demand for Personalized License Plates.” *National Tax Journal* 40 (March): 103-109.

Beaman, J., Hegmann, S., and Duwors. R.(1991). “Price Elasticity of Demand: A Campground Example.” *Journal of Travel Research* 30(1) 22/29doi: 10.1177/004728759103000105

UNIT 6 DEMAND ESTIMATION AND FORECASTING

Objectives

By studying this unit, you should be able to:

- **identify** a wide range of demand estimation and forecasting methods;
- **apply** these methods and to understand the meaning of the results;
- **understand** the nature of a demand function;
- **identify** the strengths and weaknesses of the different methods;
- **understand** that demand estimation and forecasting is about minimizing risk.

Structure

- 6.1 Introduction
- 6.2 Estimating Demand Using Regression Analysis
- 6.3 Evaluating the Accuracy of the Regression Equation - Regression Statistics
- 6.4 The Marketing Approach to Demand Measurement
- 6.5 Demand Forecasting Techniques
- 6.6 Barometric Forecasting
- 6.7 Forecasting Methods: Regression Models
- 6.8 Summary
- 6.9 Key Words
- 6.10 Self-Assessment Questions
- 6.11 Further Readings

6.1 INTRODUCTION

The first question which arises is, what is the difference between demand estimation and demand forecasting? The answer is that estimation attempts to quantify the links between the level of demand and the variables which determine it. Forecasting, on the other hand, attempts to predict the overall level of future demand rather than looking at specific linkages. For this reason, the set of techniques used may differ, although there will be some overlap between the two. In general, an estimation technique can be used to forecast demand but a forecasting technique cannot be used to estimate demand. A manager who wishes to know how high demand is likely to be in two years' time might use a forecasting technique. A manager who wishes to know how the firm's pricing policy could be used to generate a given increase in demand would use an estimation technique.

The firm needs to have information about likely future demand in order to pursue optimal pricing strategy. It can only charge a price that the market will bear if it is to sell the product. On one hand, over-optimistic estimates of demand may lead to an excessively high price and lost sales. On the other hand, over-pessimistic estimates of demand may lead to a price which is set too low resulting in lost profits. The more accurate, information the firm has, the less likely it is to take a decision which will have a negative impact on its operations and profitability.

The level of demand for a product will influence decisions, which the firm will take regarding the non-price factors that form part of its overall competitive strategy. For example, the level of advertising it carries out will be determined by the perceived need to stimulate demand for the product. As advertising expenditure represents an additional cost to the firm, unnecessary spending in this area needs to be avoided. If the firm's expectations about demand are too low it may try to compensate by spending large sums on advertising, money which in this instance may be, at least, partly wasted. Alternatively, it may decide to redesign the product in response to this, thus incurring unnecessary additional costs in the form of research and development expenditure.

In the previous unit, demand analysis was introduced as a tool for managerial decision-making. For example, it was shown that knowledge of price and cross elasticities can assist managers in pricing and that income elasticities provide useful insights into how demand for a product will respond to different macroeconomic conditions. We assumed that these elasticities were known or that the data were already available to allow them to be easily computed. Unfortunately, this is not usually the case. For many business applications, the manager who desires information about elasticities must develop a data set and use statistical methods to estimate a demand equation from which the elasticities can then be calculated. This estimated equation could then, also be used to predict demand for the product, based on assumptions about prices, income, and other factors. In this unit the basic techniques of demand estimation and forecasting are introduced.

6.2 ESTIMATING DEMAND USING REGRESSION ANALYSIS

The basic regression tools discussed in Block 1 can also be used to estimate demand relationships. Consider a small restaurant chain specializing in Chinese dinners. The business has collected information on prices and the average number of meals served per day for a random sample of eight restaurants in the chain.

These data are shown below. Use regression analysis to estimate the coefficients of the demand function $Q_d = a + bP$. Based on the estimated equation, calculate the point price elasticity of demand at mean values of the variables.

City	Meals per Day (Q)	Price (P)
1	100	150
2	90	180
3	85	190
4	110	140
5	120	130
6	90	190
7	105	160
8	100	140

Solution: The mean values of the variables are $\bar{Q} = 100$ and $\bar{P} = 160$. The other data needed to calculate the coefficients of the demand equation are shown below.

City	$Q_i - \bar{Q}$	$P_i - \bar{P}$	$(P_i - \bar{P})^2$	$(P_i - \bar{P})(Q_i - \bar{Q})$
1	0	-1	100	0
2	-10	2	400	-20
3	-15	3	900	-45
4	10	-2	400	-20
5	20	-3	900	-60
6	-10	3	900	-30
7	5	0	0	0
8	0	-2	400	0
			$\Sigma(P_i - \bar{P})^2 = 4000$	$S(P_i - \bar{P})(Q_i - \bar{Q}) = -1750$

As shown, the sum of the $(P_i - \bar{P})^2$ is 4000 and the sum of the $(P_i - \bar{P})(Q_i - \bar{Q})$ is -1750. Thus, using the equations for calculating $\hat{b}$ and $\hat{a}$, $\hat{b} = -\frac{175}{40} = -4.375$ and $\hat{a} = 100 - (4.375)(160) = 170$.

Hence, the estimated demand equation is $Q_d = 170 - 4.375P$. Recall from the previous unit that the formula for point price elasticity of demand is $E_p = (dQ/dP)(P/Q)$. Based on the estimated demand function, $dQ/dP = -4.375$. Thus, using the mean values for the price and quantity variables, $E_p = (-4.375)(160/100) = -0.7$.

6.3 EVALUATING THE ACCURACY OF THE REGRESSION EQUATION - REGRESSION STATISTICS

Once the parameters have been estimated, the strength of the relationship between the dependent variable and the independent variables can be measured in two ways. The first uses a measure called the coefficient of determination, denoted as R^2 , to measure how well the overall equation explains changes in the dependent variable. The second measure uses the t-statistic to test the strength of the relationship between an independent variable and the dependent variable.

Testing Overall Explanatory Power: Define the squared deviation of any Y_i from the mean of Y [i.e., $(Y_i - \bar{Y})^2$] as the variation in Y. The total variation is found by summing these deviations for all values of the dependent variable as total variation = $\Sigma(Y_i - \bar{Y})^2$

Total variation can be separated into two components: explained variation and unexplained variation. These concepts are explained below, for each X_i value, compute the predicted value of Y_i (denoted as $\hat{Y}_i$) by substituting X_i in the estimated regression equation:

$$\hat{Y}_i = \hat{a} + \hat{b}X_i$$

The squared difference between the predicted value $\hat{Y}_i$ and the mean value $\bar{Y}$ [i.e., $(\hat{Y}_i - \bar{Y})^2$] defined as explained variation. The word explained means that the deviation of Y from its average value is $\bar{Y}$ the result of (i.e., is

explained by) changes in X. For example, in the data on total output and cost used previously, one important reason the cost values are higher or lower than $\bar{Y}$ is because output rates (X_i) are higher or lower than the average output rate.

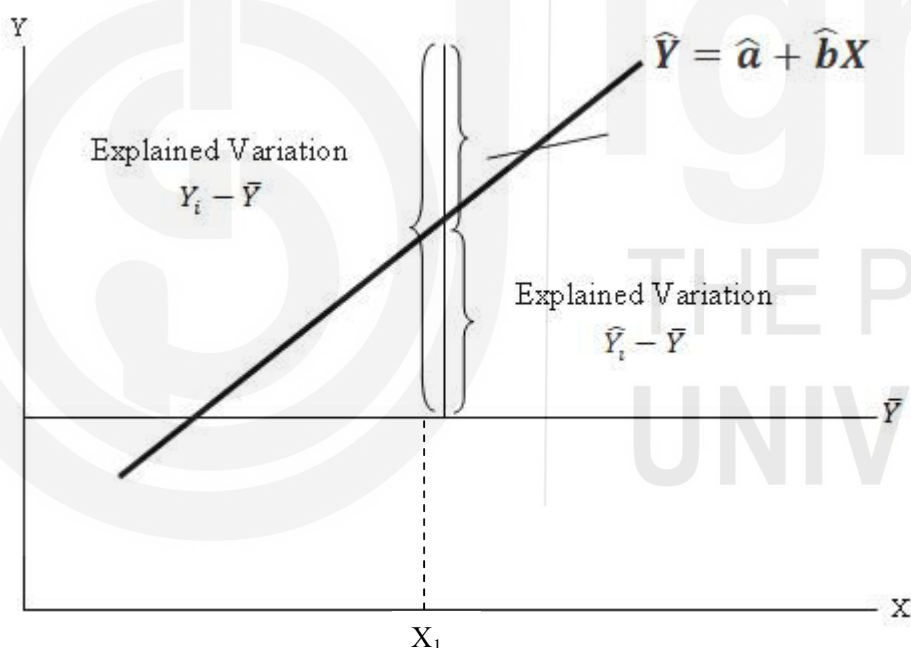
Total explained variation is found by summing these squared deviations, that is,

$$\text{total explained variation} = \sum (\hat{Y}_i - \bar{Y})^2$$

Unexplained variation is the difference between Y_i and $\hat{Y}_i$. That is, part of the deviation of Y_i from the average value ($\bar{Y}$) is "explained" by the independent variable, X. The remaining deviation, $Y_i - \hat{Y}_i$, is said to be unexplained. Summing the squares of these differences yields total unexplained variation = $\sum (Y_i - \hat{Y}_i)^2$

The three sources of variation are shown in Figure 6.1.

Figure 6.1: Sources of Variation in the Estimated Regression Equation



The coefficient of determination (R^2) measures the proportion of total variation in the dependent variable that is "explained" by the regression equation. That is,

$$R^2 = \frac{\text{total explained variation}}{\text{total variation}} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}$$

The value of R^2 ranges from zero to 1. If the regression equation explains none of the variation in Y (i.e., there is no relationship between the independent variables and the dependent variable), R^2 will be zero. If the equation explains all the variation (i.e., total explained variation = total variation), the coefficient of determination will be 1. In general, the higher the value of R^2 , the "better" the regression equation. The term fit is often

used to describe the explanatory power of the estimated equation. When R^2 is high, the equation is said to fit the data well. A low R^2 would be indicative of a rather poor fit.

Table 6.1: Computing the Sources of Variation in a Regression Mode

Y_i	Total Variation $(Y_i - \bar{Y})^2$	$\hat{Y}_i$	Explained Variation $(\hat{Y}_i - \bar{Y})^2$	Unexplained Variation $(Y_i - \hat{Y}_i)^2$
100	18,807.38	87.08	22,518.00	166.93
150	7,593.38	148.13	7,922.78	3.50
160	5,950.58	184.76	2,743.66	613.06
240	8.18	209.18	781.76	949.87
230	50.98	270.23	1,094.95	1,618.45
370	17,651.78	367.91	17,100.79	4.37
410	29,880.58	392.33	24,083.94	312.23
$Y_i=237.14$	$S(Y_i - \bar{Y})^2$ =79,942.86		$S(\hat{Y}_i - \bar{Y})^2$ =76,245.88	$S(Y_i - \hat{Y}_i)^2$ =3,668.41

How high must the coefficient of determination be in order that a regression equation be said to fit well? There is no precise answer to this question. For some relationships, such as that between consumption and income over time, one might expect R^2 to be at least 0.95. In other cases, such as estimating the relationship between output and average cost for fifty different producers during one production period, an R^2 of 0.40 or 0.50 might be regarded as quite good.

Based on the estimated regression equation for total cost and output, that is

$$\hat{Y}_1 = 87.08 + 12.21X_1$$

the coefficient of determination can be computed using the data on sources of variation shown in Table 6.1.

$$R^2 = \frac{\text{total explained variation}}{\text{total variation}} = \frac{76,245.88}{79,942.86} = 0.954$$

The value of R^2 is 0.954, which means that more than 95 percent of the variation in total cost is explained by changes in output levels. Thus, the equation would appear to fit the data quite well.

Evaluating the Explanatory Power of Individual Independent Variables

The t-test is used to determine whether there is a significant relationship between the dependent variable and each independent variable. This test requires that the standard deviation (or standard error) of the estimated regression coefficient be computed. The relationship between a dependent variable and an independent variable is not fixed because the estimate of b will vary for different data samples.

The standard error of $\hat{b}$ from one of these regression equations provides an estimate of the amount of variability in b . The equation for this standard error is

$$S_{\hat{b}} = \sqrt{\frac{\sum(Y_i - \hat{Y}_i)^2 / (n - 2)}{\sum(X_i - \bar{X})^2}}$$

where n is the number of observations. For the production-cost example used in this section, $n = 7$ and the standard error of $\hat{b}$ is

$$S_{\hat{b}} = \sqrt{\frac{3,668.41/5}{511.40}} = 1.19$$

The least-squares estimate of $\hat{b}$ is said to be an estimate of the parameter b . But it is known that $\hat{b}$ is subject to error and thus will differ from the true value of the parameter b . That is why $\hat{b}$ is called an estimate.

Because of the variability in $\hat{b}$, it sometimes is useful to determine a range or interval for the estimate of the true parameter b . Using principles of statistics, a 95 percent confidence interval estimate for b is given by the equation

$$\hat{b} \pm t_{n-k-1} S_{\hat{b}}$$

where t_{n-k-1} represents the value of a particular probability distribution known as student's distribution. The subscript $(n - k - 1)$ refers to the number of degrees of freedom, where n is the number of observations or data points and k is the number of independent variables in the equation. An abbreviated list of t -values for use in estimating 95 percent confidence intervals is shown in Table 6.4. In the example discussed here, $n = 7$ and $k = 1$, so there are five (i.e., $7 - 1 - 1$) degrees of freedom, and the value of t in the table is 2.571. Thus, in repeated estimations of the output cost relationship, it is expected that about 95 percent of the time such that the true value of parameter b will lie in the interval defined by the estimated value of b plus or minus 2.571 times the standard error of b . For output-cost data, the 95 percent confidence interval estimate would be

$$12.21 \pm 2.571(1.19)$$

or from 9.15 to 15.27. This means that the probability that the true marginal relationship between cost and output (i.e., the value of b) within this range is 0.95.

If there is no relationship between the dependent and an independent variable, the parameter b would be zero. A standard statistical test for the strength of the relationship between Y and X is to check whether the 95 percent confidence interval includes the value zero. If it does not, the relationship between X and Y as measured by $\hat{b}$ is said to be statistically significant. If that interval does include zero, then b is said to be non-significant, meaning that there does not appear to be a strong relationship between the two variables. The confidence interval for $\hat{b}$ in the output-cost example did not include zero, and thus it is said that $\hat{b}$, an estimate of marginal cost, is statistically significant or that there is a strong relationship between cost and rate of output.

Another way to make the same test is to divide the estimated coefficient ($\hat{b}$) by its standard error. The probability distribution of this ratio is the same as Student's t distribution; thus, this ratio is called a t -value. If the absolute value of this ratio is equal to or greater than the tabled value of t for $n - k - 1$

degrees of freedom, $\hat{b}$ is said to be statistically significant. Using the output-cost data, the t-value is computed to be

$$t = \left| \frac{\hat{b}}{s_{\hat{b}}} \right| = \left| \frac{12.21}{1.19} \right| = 10.26$$

Because the ratio is greater than 2.571, the value of the t-statistic from Table 6.2, it is concluded that there is a statistically significant relationship between cost and output.

In general, if the absolute value of the ratio $\hat{b} / S_{\hat{b}}$ is greater than the value from the table for $n - k - 1$ degrees of freedom, the coefficient $\hat{b}$ is said to be statistically significant.

Table 6.2: Selected Values of the Student's Distribution for 95 Per cent Confidence Interval

Degree of Freedom	t-value
1	12.706
3	3.182
5	2.571
7	2.365
10	2.228
20	2.086
30	2.043
60	2.000
120	1.980

The standard error of the equation is used to determine the likely accuracy with which we can predict the value of the dependent variable associated with particular values of the independent variables. As a general principle, the smaller the value of the standard error of the equation, the more accurate the equation is and hence the more accurate any predictions made from it will be. To put this in another way, the standard error represents the standard deviation of the dependent variable about the regression line. Thus, the smaller the value, the better the fit of the equation to the data and the closer the estimate will be to the true regression line. Conversely, the larger the standard error, the bigger the deviation from the regression line and the less confidence that can be put in any prediction arising from it. The standard error of the coefficient works along similar lines. It gives an indication of the amount of confidence that can be placed in the estimated regression coefficient for each independent variable. Again, the smaller the value, the greater the confidence that can be placed in the estimated coefficient and vice versa. Finally, the t-test provides a further measurement of the accuracy of the regression coefficient for each of the independent variables.

A value of t greater than or equal to 2 generally indicates that the calculated coefficient is a reliable estimate, while a value of less than 2 indicates that the coefficient is unreliable.

(Note: This also partly depends, however, on the number of data observations on which the equation is based so that t-test tables need to be used in order to ensure an accurate interpretation of this statistic.)

Activity 1

Having described the statistics let us now consider how they may be used in practice. To do this, we can add example regression statistics to the previously estimated regression equation.

$$\log QD = \log 200 - 1.5 \log P_o + 2.4 \log A_o$$

(0.3) (0.4)

standard errors of the respective estimates are in parenthesis

$$R^2 = 0.95, k = 2 \text{ and } n = 20$$

What does the R^2 tell us? Are both the coefficients reliable? Explain.

.....

.....

.....

.....

6.4 THE MARKETING APPROACH TO DEMAND MEASUREMENT

The vast majority of business decisions involve some degree of uncertainty and managers seldom know exactly what the outcomes of their choices will be. One approach to reducing the uncertainty associated with decision making is to devote resources to forecasting. Forecasting involves predicting future economic conditions and assessing their effect on the operations of the firm.

Frequently, the objective of forecasting is to predict demand. In some cases, managers are interested in the total demand for a product. For example, the decision by an office products firm to enter the home computer market may be determined by estimates of industry sales growth. In other circumstances, the projection may focus on the firm's probable market share. If a forecast suggests that sales growth by existing firms will make successful entry unlikely, the company may decide to look for other areas in which to expand.

Forecasts can also provide information on the proper product mix. For an automobile manufacturer such as Maruti Suzuki Ltd., managers must determine the number of Swift versus Alto to be produced. In the short run, this decision is largely constrained by the firm's existing production facilities for producing each kind of car. However, over a longer period, managers can build or modify production facilities. But such choices must be made long before the vehicles begin coming off the assembly line. Accurate forecasts can reduce the uncertainty caused by this long lead time. For example, if the price of petrol is expected to increase, the relative demand for Alto or compact cars is also likely to increase.

Forecasting is an important management activity. Major decisions in large businesses are almost always based on forecasts of some type. In some cases, the forecast may be little more than an intuitive assessment of the future by

those involved in the decision. In other circumstances, the forecast may have required thousands of work hours and lakhs of rupees. It may have been generated by the firm's own economists, provided by consultants specializing in forecasting, or be based on information provided by government agencies. Forecasting requires the development of a good set of data on which to base the analysis. A forecast cannot be better than the data from which it is derived. Three important sources of data used in forecasting are expert opinion, surveys, and market experiments.

Expert Opinion

The collective judgment of knowledgeable persons can be an important source of information. In fact, some forecasts are made almost entirely on the basis of the personal insights of key decision makers. This process may involve managers conferring to develop projections based on their assessment of the economic conditions facing the firm. In other circumstances, the company's sales personnel may be asked to evaluate future prospects. In still other cases, consultants may be employed to develop forecasts based on their knowledge of the industry. Although predictions by experts are not always the product of "hard data," their usefulness should not be underestimated. Indeed, the insights of those closely connected with an industry can be of great value in forecasting.

Methods exist for enhancing the value of information elicited from experts. One of the most useful is the **Delphi technique**. Its use can be illustrated by a simple example. Suppose that a panel of six outside experts is asked to forecast a firm's sales for the next year. Working independently, two panel members forecast an 8 percent increase, three members predict a 5 percent increase, and one person predicts no increase in sales. Based on the responses of the other individuals, each expert is then asked to make a revised sales forecast. Some of those expecting rapid sales growth may, based on the judgments of their peers, present less optimistic forecasts in the second iteration. Conversely, some of those predicting slow growth may adjust their responses upward. However, there may also be some panel members who decide that no adjustment of their initial forecast is warranted.

Assume that a second set of predictions by the panel includes one estimate of a 2 percent sales increase, one of 5 percent, two of 6 percent, and two of 7 percent. The experts again are shown each other's responses and asked to consider their forecasts further. This process continues until a consensus is reached or until further iterations generate little or no change in sales estimates.

The value of the Delphi technique is that it aids individual panel members in assessing their forecasts. Implicitly, they are forced to consider why their judgment differs from that of other experts. Ideally, this evaluation process should generate more precise forecasts with each iteration.

One problem with the Delphi method can be its expense. The usefulness of expert opinion depends on the skill and insight of the experts employed to make predictions. Frequently, the most knowledgeable people in an industry are in a position to command large fees for their work as consultants or they may be employed by the firm, but have other important responsibilities,

which means that there can be a significant opportunity cost in involving them in the planning process. Another potential problem is that those who consider themselves experts may be unwilling to be influenced by the predictions of others on the panel. As a result, there may be few changes in subsequent rounds of forecasts.

Surveys

Surveys of managerial plans can be an important source of data for forecasting. The rationale for conducting such surveys is that plans generally form the basis for future actions. For example, capital expenditure budgets for large corporations are usually planned well in advance. Thus, a survey of investment plans by such corporations should provide a reasonably accurate forecast of future demand for capital goods.

Several private and government organizations conduct periodic surveys. The annual National Council of Applied Economic Research (NCAER) survey of Market Information of Households is well recognized. Many private organizations like ORG-MARG and TNS INDIA conduct surveys relating to consumer demand across certain geographical areas.

If data from existing sources do not meet its specific needs, a firm may conduct its own survey. Perhaps the most common example involves companies that are considering a new product or making a substantial change in an existing product. But with new or modified products, there are no data on which to base a forecast. One possibility is to survey households regarding their anticipated demand for the product. Typically, such surveys attempt to ascertain the demographic characteristics (e.g., age, education, and income) of those who are most likely to buy the product and find how their decisions would be affected by different pricing policies.

Although surveys of consumer demand can provide useful data for forecasting, their value is highly dependent on the skills of their originators. Meaningful surveys require careful attention to each phase of the process. Questions must be precisely worded to avoid ambiguity. The survey sample must be properly selected so that responses will be representative of all customers. Finally, the methods of survey administration should produce a high response rate and avoid biasing the answers of those surveyed. Poorly phrased questions or a non random sample may result in data that are of little value.

Even the most carefully designed surveys do not always predict consumer demand with great accuracy. In some cases, respondents do not have enough information to determine if they would purchase a product. In other situations, those surveyed may be pressed for time and be unwilling to devote much thought to their answers. Sometimes the response may reflect a desire (either conscious or unconscious) to put oneself in a favorable light or to gain approval from those conducting the survey. Because of these limitations,

forecasts seldom rely entirely on results of consumer surveys. Rather, these data are considered supplemental sources of information for decision making.

Market Experiments

A potential problem with survey data is that survey responses may not translate into actual consumer behavior. That is, consumers do not necessarily do what they say they are going to do. This weakness can be partially overcome by the use of market experiments designed to generate data prior to the full-scale introduction of a product or implementation of a policy.

To set up a market experiment, the firm first selects a test market. This market may consist of several cities; a region of the country, or a sample of consumers taken from a mailing list. Once the market has been selected, the experiment may incorporate a number of features. It may involve evaluating consumer perceptions of a new product in the test market. In other cases, different prices for an existing product might be set in various cities in order to determine demand elasticity. A third possibility would be a test of consumer reaction to a new advertising campaign.

There are several factors that managers should consider in selecting a test market. First, the location should be of manageable size. If the area is too large, it may be expensive and difficult to conduct the experiment and to analyze the data. Second, the residents of the test market should resemble the overall population of India in age, education, and income. If not, the results may not be applicable to other areas. Finally, it should be possible to purchase advertising that is directed only to those who are being tested.

Market experiments have an advantage over surveys in that they reflect actual consumer behavior, but they still have limitations. One problem is the risk involved. In test markets where prices are increased, consumers may switch to products of competitors. Once the experiment has ended and the price reduced to its original level, it may be difficult to regain those customers. Another problem is that the firm cannot control all the factors that affect demand. The results of some market experiments can be influenced by bad weather, changing economic conditions, or the tactics of competitors. Finally, because most experiments are of relatively short duration, consumers may not be completely aware of pricing or advertising changes. Thus, their responses may understate the probable impact of those changes.

Activity 2

What are the major marketing approaches to demand measurement?

.....

.....

.....

6.5 DEMAND FORECASTING TECHNIQUES

Time-series analysis

Regression analysis, as described above, can be used to quantify relationships between variables. However, data collection can be a problem if the regression model includes a large number of independent variables. When changes in a variable show discernable patterns over time, time-series analysis is an alternative method for forecasting future values.

The focus of time-series analysis is to identify the components of change in the data. Traditionally, these components are divided into four categories:

1. Trend
2. Seasonality
3. Cyclical patterns
4. Random fluctuations

A **trend** is a long-term increase or decrease in the variable. For example, the time series of population in India exhibits an upward trend, while the trend for endangered species, such as the tiger, is downward. The seasonal component represents changes that occur at regular intervals. A large increase in sales of umbrellas during the monsoon would be an example of seasonality.

Analysis of a time series may suggest that there are **cyclical patterns**, defined as sustained periods of high values followed by low values. Business cycles fit this category. Finally, the remaining variation in a variable that does not follow any discernable pattern is due to **random fluctuations**. Various methods can be used to determine trends, seasonality, and any cyclical patterns in time-series data. However, by definition, changes in the variable due to random factors are not predictable. The larger the random component of a time series, the less accurate the forecasts based on those data.

Trend Projection

One of the most commonly used forecasting techniques is trend projection. As the name suggests, this approach is based on the assumption that there is an identifiable trend in a time series of data. Trend projection can also be used as the starting point for identifying seasonal and cyclical variations.

Table 6.3 is a time series of a firm's quarterly sales over a three-year time span. These data are used to illustrate graphical and statistical trend projection and also to describe a method for making seasonal adjustments to a forecast.

Table 6.3: Hypothetical Time-Series Sales Data

Period Number	Quarter	Sales (Lakhs)
1	2018:I	Rs. 300
2	2018:II	305
3	2018:III	315
4	2018:IV	340
5	2019:I	346
6	2019:II	352
7	2019:III	364
8	2019:IV	390
9	2020:I	397
10	2020:II	404
11	2020:III	418
12	2020:IV	445

Statistical Curve Fitting Basically, this involves using the ordinary least-squares concept developed above to estimate the parameters of the equation. Suppose that an analyst determines that a forecast will be made assuming that there will be a constant rate of change in sales from one period to the next. That is, the firm's sales will change by the same amount between two periods. The time-series data of Table 6.4 are to be used to estimate that rate of change.

Statistically, this involves estimating the parameters of the equation

$$S_t = S_0 + bt$$

where S denotes sales and t indicates the time period. The two parameters to be estimated are S_0 and b. The value of S_0 corresponds vertical intercept of the line and the parameter b is the constant rate of change and corresponds to the slope. Many hand calculators can estimate the parameters of equation. Specific procedures vary from model to model, but usually the only requirement is that the users input the data and push one or two designated keys. The machine then returns the estimated parameters. For the data of Table 6.3, the quarters would have to be input as sequential numbers starting with 1. That is, 2018: I would be entered as 1, 2018: II would be entered as 2, and so forth. Based on the data from the table, the equation is estimated as

$$S_f = 281.394 + 12.811t$$

The interpretation of the equation is that the estimated constant rate of increase in sales per quarter is Rs. 12.811 lakhs. A forecast of sales for any future quarter, S_t , can be obtained by substituting in the appropriate value for t. For example, the third quarter of 2021 is the 15th observation of the time series. Thus, the estimated sales for that quarter would be $281.394 + 12.811(15)$, or Rs. 473.56 lakhs.

Now suppose that the individual responsible for the forecast wants to estimate a percentage rate of change in sales. That is, it is assumed that sales will increase by a constant percent each period. This relationship can be expressed mathematically as

$$S_t = S_{t-1}(1 + g)$$

Similarly,

$$S_{t-1} = S_{t-2}(1 + g)$$

where g is the constant percentage rate of change, or the growth rate. These two equations imply that

$$S_t = S_{t-2} (1 + g)^2$$

and, in general, $S_t = S_0 (1 + g)^t$

As shown, the parameters of this equation cannot be estimated using ordinary least squares. The problem is that the equation is not linear. However, there is a simple transformation of the equation that allows it to be estimated using ordinary least squares.

Take logs, the result is

$$\ln S_t = \ln [S_0 (1 + g)^t]$$

But the logarithm of a product is just the sum of the logarithms. Thus

$$\ln S_t = \ln S_0 + \ln[(1 + g)^t]$$

The right-hand side of the equation can be further simplified by noting that

$$\ln [(1 + g)^t] = t[\ln(1 + g)]$$

Hence

$$\ln S_t = \ln S_0 + t(\ln(1 + g))$$

This equation is linear in form. This can be seen by making the following substitutions:

$$Y_t = \ln S_t$$

$$Y_0 = \ln S_0$$

$$b = \ln(1 + g)$$

Thus the new equation is

$$Y_t = Y_0 + bt$$

which is linear.

The parameters of this equation can easily be estimated using a hand calculator. The key is to recognize that the sales data have been translated into logarithms. Thus, instead of S_t , it is $\ln S_t$ that must be entered as data. However, note that the t values have not been transformed. Hence, for the first quarter of 1996, the data to be entered are $\ln 300 = 5.704$ and 1; for the second quarter, $\ln 305 = 5.720$ and 2; and so forth. The transformed data are provided in Table 6.4

Table 6.4: Natural Logarithms of Hypothetical Time-Series Sales Data

Period Number (t)	Quarter	Natural Logarithm of Sales (in Lakhs) S_t
1	2018:I	5.704
2	2018:II	5.720
3	2018:III	5.753
4	2018:IV	5.829
5	2019:I	5.847
6	2019:II	5.864
7	2019:III	5.897
8	2019:IV	5.966
9	2020:I	5.984
10	2020:II	6.001
11	2020:III	6.036
12	2020:IV	6.098

Using the ordinary least-squares method, the estimated parameters of the equation based on the data from Table 6.5 are

$$Y_t = 5.6623 + 0.0353t$$

But these parameters are generated from the logarithms of the data. Thus, for interpretation in terms of the original data, they must be converted back based on the relationships $\ln S_0 = Y_0 = 5.6623$ and $\ln(1 + g) = b = 0.0353$. Taking the antilogs yields $S_0 = 287.810$ and $1 + g = 1.0359$. Substituting these values for S_0 and $1 + g$ back into the original equation gives

$$S_t = 287.810(1.0359)^t$$

where 287.810 is sales (in lakhs of rupees) in period 0 and the estimated growth rate, g , is 0.0359 or 3.59 per cent.

To forecast sales in a future quarter, the appropriate value of t is substituted into the equation. For example, predicted sales in the third quarter of 2021 (i.e., the fifteenth quarter) would be $287.810(1.0359)^{15}$, or ₹ 488.51 lakhs.

Seasonal Variation in Time-Series Data

Seasonal fluctuations in time-series data are not uncommon. In particular, a large increase in sales for the fourth quarter is a characteristic of certain industries. Indeed, some retailing firms make large amounts of their total sales during the Diwali period. Other business activities have their own seasonal sales patterns. Electric companies serving hot, humid areas have distinct peak sales periods during the summer months because of the extensive use of air conditioning. Similarly, demand for accountants' services increases in the first quarter as income tax deadlines approach.

A close examination of the data in Table 6.4 indicates that the quarterly sales increases are not uniformly distributed over the year. The increases from the first quarter to the second, and from the fourth quarter to the first, tend to be small, while the fourth-quarter increase is consistently larger than that of other quarters. That is, the data exhibits seasonal fluctuations.

Pronounced seasonal variations can cause serious errors in forecasts based on time-series data. For example, Table 6.4 indicates that actual sales for the fourth quarter 2020 were ₹ 445 lakhs. But if the estimated equation is used to predict sales for that period (using the constant rate of change model), the predicted total is $281.394 + 12.811(12)$, or ₹ 435.13 lakhs. The large difference between actual and predicted sales occurs because the equation does not take into account the fourth- quarter sales jump. Rather, the predicted value from the equation represents an averaging of individual quarters. Thus, sales will be underestimated for the strong fourth quarter. Conversely the predicting equation may overestimate sales for other quarters.

The accuracy of the forecast can be improved by seasonally adjusting the data. Probably the most common method of adjustment is the ratio-to-trend approach. Its use can be illustrated using the data from Table 6.4 based on predicting equation,

$$S_t = 281.394 + 12.811t$$

actual and calculated fourth-quarter sales are shown in Table 6.5. The final column of the table is the ratio of actual to predicted sales for the fourth quarter. This ratio is a measure of the seasonal error in the forecast.

As shown, for the three-year period, average actual sales for the fourth quarter were 102 percent of the average forecasted sales for that quarter. The factor 1.02 can be used to adjust future fourth-quarter sales estimates. For example, if the objective is to predict sales for the fourth quarter of 2020, the predicting equation generates an estimate of Rs. 435.13 lakhs. Multiplying this number by the 1.020 adjustment factor, the forecast is increased to Rs. 443.8 lakhs, which is close to the actual sales of Rs. 445 lakhs for that quarter. A similar technique could be used to make a downward adjustment for predicted sales in other quarters.

Seasonal adjustment can improve forecasts based on trend projection. However, trend projection still has some shortcomings. One is that it is primarily limited to short-term predictions. If the trend is extrapolated much beyond the last data point,

Table 6.5: Seasonal Adjustment Using the Ratio-to-Trend Method

Year	Forecasted Fourth-Quarter Sales	Actual Fourth-Quarter Sales	Actual/Predicted Fourth-Quarter Sales
2018	332.64	₹340	1.022
2019	383.88	390	1.016
2020	435.13	445	1.023
		Average =	1.020

the accuracy of the forecast diminishes rapidly. Another limitation is that factors such as changes in relative prices and fluctuations in the rate of economic growth are not considered. Rather, the trend projection approach assumes that historical relationships will not change.

Exponential Smoothing

Trend projection is actually just regression analysis where the only independent variable is time. One characteristic of this method is that each observation has the same weight. That is, the effect of the initial data point on the estimated coefficients is just as great as the last data point. If there has been little or no change in the pattern over the entire time series, this is not a problem. However, in some cases, more recent observations will contain more accurate information about the future than those at the beginning of the series. For example, the sales history of the last three months may be more relevant in forecasting future sales than data for sales 10 years in the past.

Exponential smoothing is a technique of time-series forecasting that gives greater weight to more recent observations. The first step is to choose a smoothing constant, α , where $0 < \alpha < 1.0$. If there are n observations in a time series, the forecast for the next period (i.e., $n + 1$) is calculated as a weighted average of the observed value of the series at period n and the forecasted value for that same period. That is,

$$F_{n+1} = \alpha X_n + (1 - \alpha)F_n$$

where F_{n+1} is the forecast value for the next period, X_n is the observed value for the last observation, and F_n is a forecast of the value for the last period in the time series. The forecasted values for F_n and all the earlier periods are calculated in the same manner.

Specifically,

$$F_t = \alpha X_{t-1} + (1 - \alpha)F_{t-1}$$

starting with the second observation (i.e., $t = 2$) and going to the last (i.e., $t = n$). Note that equation cannot be used to forecast F_1 because there is no X_0 or F_0 . This problem is usually solved by assuming that the forecast for the first period is equal to the observed value for that period. That is, $F_1 = X_1$. Using the equation it can be seen that this implies that the second-period forecast is just the observed value for the first period, or $F_2 = X_1$.

The exponential smoothing constant chosen determines the weight that is given to different observations in the time series. As α approaches 1.0, more recent observations are given greater weight. For example, if $\alpha = 1.0$, then $(1 - \alpha) = 0$ and the equations indicate that the forecast is determined only by the actual observation for the last period. In contrast, lower values for α give greater weight to observations from previous periods.

Table 6.6 Forecasts Based on Exponential Smoothing

Week(T)	Sales(X_t)	$a=0.20(F_t)$	$a=0.40(F_t)$	$a=0.60(F_t)$	$a=0.80(F_t)$
1	400	400.00	400.00	400.00	400.00
2	430	400.00	400.00	400.00	400.00
3	420	406.00	412.00	418.00	424.00
4	440	408.80	415.20	419.20	420.80
5	460	415.04	425.12	431.68	436.18
6	440	424.03	439.07	448.67	455.23
7	470	427.23	439.44	443.47	443.05
8	430	435.78	451.67	459.39	464.61
9	440	434.62	443.00	441.76	436.92
10	420	435.70	441.80	440.70	439.38
11		432.56	433.08	428.28	423.88

Assume that a firm's sales over the last 10 weeks are as shown in Table 6.6.

By assumption, $F_2 = F_1 = X_1$ if $\alpha = 0.20$, then

$$F_3 = 0.20(4.30) + 0.80(400) = 406.0$$

and

$$F_4 = 0.20(420) + 0.80(406) = 408.8$$

The forecasted values for four different values of a are provided in Table 6.6.

The table also shows forecasted sales for the next period after the end of the time-series data, or week 11. Using $a = 0.20$, the forecasted sales value for the 11th week is computed to be

$$F_{11} = 0.20(420) + 0.80(435.7) = 432.56$$

Table 6.6 suggests why this method is referred to as smoothing technique. Consider the forecasts based on $a = 0.20$. Note that the smoothed data show much less fluctuation than the original sales data. Note also that as a increases, the fluctuations in the F_t increase, because the forecasts give more weight to the last observed value in the time series.

Choice of a Smoothing Constant

Any value of α could be used as the smoothing constant. One criterion for selecting this value might be the analyst's intuitive judgment regarding the weight that should be given to more recent data points. But there is also an empirical basis for selecting the value of α . Remember that the coefficients of a regression equation are chosen to minimize the sum of squared deviations between observed and predicted values. This same method can be used to determine the smoothing constant.

The term $(X_t - F_t)^2$ is the square of the deviation between the actual time-series data and the forecast for the same period. Thus, by adding these values for each observation, the sum of the squared deviations can be computed as

$$\sum_{t=1}^n (X_t - F_t)^2$$

One approach to choosing α is to select the value that minimizes this sum. For the data and values shown in Table 6.6, these sums are

Smoothing Constant	Sum of Squared Deviations
0.20	6484.23
0.40	4683.87
0.60	4213.08
0.80	4394.52

These results suggest that, of the four values of the smoothing constant, $\alpha = 0.60$ provides the best forecasts using these data. However, it should be noted that there may be values of α between 0.60 and 0.80 or between 0.40 and 0.60 that yield even better results.

Evaluation of Exponential Smoothing

One advantage of exponential smoothing is that it allows more recent data to be given greater weight in analyzing time-series data. Another is that, as additional observations become available, it is easy to update the forecasts. There is no need to re-estimate the equations, as would be required with trend projection.

The primary disadvantage of exponential smoothing is that it does not provide very accurate forecasts if there is a significant trend in the data. If the time trend is positive, forecasts based on exponential smoothing will be likely to be too low, while a negative time trend will result in estimates that are too high. Simple exponential smoothing works best when there is no discernable time trend in the data. There are, however, more sophisticated forms of exponential smoothing that allow both trends and seasonality to be accounted for in making forecasts.

6.6 BAROMETRIC FORECASTING

Barometric forecasting is based on the observed relationships between different economic indicators. It is used to give the decision maker an insight into the direction of likely future demand changes, although it cannot usually be used to quantify them.

Five different types of indicators may be used. Firstly, there are **leading indicators** which run in advance of changes in demand for a particular product. An example of these might be an increase in the number of building permits granted which would lead to an increase in demand for building-related products such as wood, concrete and so on. Secondly, there are **coincident indicators** which occur alongside changes in demand. Retail sales would fall into this category, as an increase in sales would generate an increase in demand for the manufacturers of the goods concerned. Thirdly, there are **lagging indicators** which run behind changes in demand. New industrial investment by firms is often said to fall into this category. In this case it is argued that firms will only invest in new production facilities when demand is already firmly established. Thus, increased investment is a sign, or confirmation, that an initial increase in demand has already taken place. This

may well indicate that the economy is improving, for example, so that further changes in the level of demand can be expected in the near future.

One particular problem with each of these three types of indicators is that single indicator does not always prove to be accurate in predicting changes in demand. For this reason, groups of indicators may be used instead. The fourth and fifth types of indicators fall into this category. These are **composite indices** and **diffusion indices** respectively. Composite indices are made up of weighted averages of several leading indicators which demonstrate an overall trend. Diffusion indices are groups of leading indicators whose directional shifts are analysed separately. If more than half of the leading indicators included within them are rising, demand is forecast to rise and vice versa. Again, it is important to note that it is the direction of change that is the basis of the prediction, the actual size of the change cannot be measured. In addition, the situation is complicated by the fact that there may be variations in the length of the lead time between the various indicators. This means that the accuracy of predictions may be reduced.

6.7 FORECASTING METHODS: REGRESSION MODELS

You have seen how regression analysis is used in the estimating process. In this part you will see several applications of multiple regression analysis to the forecasting process. In this section we shall forecast demand by using data for Big Sky Foods (BSF) a company selling groceries.

Using the OLS method of estimation available in Excel or any standard statistical package, the demand function we estimated was

$$Q = 15.939 - 9.057P + .009INC + 5.092PC$$

where Q = sales; P = BSF's price; INC= income; PC = price charged by BSF's major competitor. This model can be used to forecast sales, assuming that forecasts of the independent variables are available.

Table 6.7: Data Used to Estimate Big Sky Foods' Demand Function

Observation	Sales (thousands of units)	Price	Income (₹ Lakhs)	Competitor's Price
2016Q4	20	5	2620	5
2017Q1	16	5.2	2733	4.8
2017Q2	16	5.32	2898	4.8
2017Q3	14	5.48	3056	4.5
2017Q4	16	5.6	3271	4.44
2018Q1	19	5.8	3479	4.55
2018Q2	17	6.03	3736	4.6
2018Q3	18	6.01	3868	4.85
2018Q4	21	5.92	4016	5.1
2019Q1	26	5.9	4152	5.4
2019Q2	30	5.85	4336	5
2019Q3	26	5.8	4477	4.95
2019Q4	27	5.85	4619	5
2020Q1	29	5.8	4764	5

SUMMARY OUTPUT (From Excel Spreadsheet)

Regression Statistics	
R Square	0.897591549
Adjusted R Square	0.866869014
Standard Error	1.985445508
Observations	14

	Coefficients	Standard Error	t Stat
Intercept	15.9392633	20.76872418	0.76746473
Price	-9.05740835	3.119255598	-2.903708294
Income (₹ Lakhs)	0.00857807	0.001545678	5.549714378
Competitor's Price	5.09168419	2.581615572	1.972285976

Big Sky Foods has access to forecasts from one of the macro econometric service firms that provide a good estimate of the income variable by quarter for one year ahead. In addition, BSF has had reasonable success using a simple exponential smoothing model (with $w = .8$) to predict the competitor's price one quarter in advance. And, of course, BSF controls its own price.

Assume that BSF plans to price at 5.85 next quarter, that the competitor's price is forecast to be 4.99, and that income is forecast to be 4800. Sales for BSF can then be forecast as follows:

$$Q = 15.939 - 9.057(5.85) + .009(4800) + 5.092(4.99)$$

$$Q = 31.565$$

Notice that, in making this forecast, BSF starts with an economic forecast that provides a projection for income and an exponential smoothing model that provides a projected value for the competitor's price. These are then combined with the multiple regression model of demand and BSF's own pricing plan to arrive at a forecast for sales. BSF can then use this procedure to experiment with the effect of different prices or to make forecasts based on differing forecasts of the other independent variables.

Activity 3

Try this yourself. Suppose that forecasts for income and the competitor's price are the same as those in our example, and that you want to evaluate the effect of setting BSF's price at ₹5.75 rather than ₹5.85. What estimate for sales (Q) would you obtain?

$$Q = 15.939 - 9.057() + .009() + 5.092()$$

$$Q = \underline{\hspace{2cm}}$$

What can you say about price elasticity based on this result?

6.8 SUMMARY

In this chapter we have looked at a range of demand estimation and forecasting techniques which can be used by the firm either singly or in combination in order to predict the level of demand for their product(s). The

choice of technique will depend upon the resources at the firm's disposal, the cost to the firm of insufficient knowledge of the market(s) in which it operates and the ease with which information can be obtained. Each of the methods we have considered has its own advantages and disadvantages in its use and there is no 'right' or 'wrong' approach in any given situation. It is for the decision maker to choose the technique(s) which are most appropriate to the firm's needs. As a general principle, however, the more, and the more accurate, information the firm has the better able it will be to take the best decisions possible for the firm's efficient operation. Thus, the firm can substantially reduce the risk to which it will be exposed, particularly in rapidly changing markets.

Sales forecasts can be developed using qualitative methods, such as expert opinion, the Delphi method, or market surveys or by using quantitative models, such as exponential smoothing, time series decomposition, or multiple regression analysis. In many cases, firms use a combination of qualitative and quantitative forecasting techniques. The use of more than one sales forecast method is advisable because doing so can reduce errors in the final forecast.

6.9 KEY WORDS

Demand forecasting is a prediction or estimation of the future demand.

Regression makes use of both economic theory and estimation techniques to generate forecasts from historical data.

Trend Method is a forecasting technique, where the time series data on the variable under forecast are used to fit a trend line or curve either graphically or by means of a statistical technique known as the Least-Squares method.

6.10 SELF-ASSESSMENT QUESTIONS

Look at Table 6.7 in this unit. That table contains a set of data related to Big Sky Foods' sales and price, consumer income, and the price charged by their major competitor. The data cover the period 2016 Q_4 – 2020 Q_1 . As in the previous problem, you should ignore the actual values given for the first quarter of 2020 and see how well you can forecast them using the tools covered in this chapter.

- Start by estimating a new demand function using just the first 13 observations, with sales (S) a function of price (P), income (INC), and the competitor's price (CP). Write your function and related statistical results here:

$$S = a + b_1P + b_2INC + b_3PC$$

() () () Put t-ratios in the parentheses.

$$R^2 =$$

- b. Now, estimate a simple linear time trend for income based on data for 2016 Q_4 – 2019 Q_4 :

$$INC = a + bT$$

() t-ratio

Project the trend ahead one quarter to forecast income for the first quarter of 1995: Income forecast for 2020 Q_1 =

- c. Use an exponential smoothing model to make a forecast of the competitor's price (CP) for the first quarter of 2020:

Competitor's price forecast for 2020 Q_1 =

- d. Assuming that Big Sky Foods does intend to set its price at ₹5.80 during the first quarter of 2017, use the information in parts a through c to make a sales forecast for 2020 Q_1 .

First 2020 Q_1 sales forecast =

- e. Now, prepare another sales forecast based on just a simple linear time trend of the sales data:

Second 2020 Q_1 sales forecast =

- f. Given that the actual level of sales in the first quarter of 2017 was 29, which model gave the best forecast? Without knowing actual sales, how might you have judged the two models used, and in which one do you think you would have had the most confidence? Why?

6.11 FURTHER READINGS

Maddala, G.S., and Ellen Miller, 1989. *Micro Economics: Theory and Applications* McGraw-Hill, New York.

Lewis, W.C., Jain, S.K., & Petersen, H.C. (2005). *Managerial Economics* (4th ed.). Pearson.

Mote, V.L., Paul, S. & Gupta, G.S. (2016). *Managerial Economics: Concepts and Cases*, Tata McGraw-Hill, New Delhi.

Dholakia, R., & Oza, A.N. (1996). *Microeconomics for Management Students*. Oxford University Press, Delhi.

Block 3

PRODUCTION AND COST ANALYSIS

Unit 7: Production Function	135
Unit 8: Short Run Cost Analysis	166
Unit 9: Long Run Cost Analysis	193

BLOCK 3 PRODUCTION AND COST ANALYSIS

Block 3 introduces production and cost analysis and the estimation of production and cost functions. Production is the process of combining inputs to create output which the firm sells in the market. The relationship between the production function and the cost function is analyzed in this block and the implication for managerial decisions is examined. In **Unit 7**, the production function and its building blocks are analyzed from a managerial perspective. The difference between short run and long run is explained from an economist's perspective. It is discussed that a profit-maximizing firm will choose an optimal combination of inputs based on input prices. Estimation of production function and its managerial uses are also elaborated in this unit. **Unit 8 and 9** examines cost analysis in the short and long run. Cost is related to production because if the production function exhibits increasing returns to scale, the unit cost of production falls, while if the production function exhibits decreasing returns to scale the unit cost will rise and so on. Different types of costs including accounting costs & opportunity costs and short run cost function are defined in **Unit 8**. Further, certain applications of short run cost function like break-even analysis and operating leverage are also explained. In **Unit 9**, long run cost function, economies of scale & scope are explained. Finally, estimation of cost function is presented using the tools of regression analysis that were developed in **Blocks 1 and 2**.

THE PEOPLE'S
UNIVERSITY

UNIT 7 PRODUCTION FUNCTION

Objectives

After going through this unit, you should be able to:

- **familiarize** with the concepts and rules relevant for production decision analysis;
- **understand** the economics of production;
- **understand** the set of conditions required for efficient production.

Structure

- 7.1 Introduction
- 7.2 Production Function
- 7.3 Production Function with one Variable inputs
- 7.4 Production Function with two Variable inputs
- 7.5 The Optimal Combination of inputs
- 7.6 Returns to Scale
- 7.7 Functional Forms of Production Function
- 7.8 Managerial Uses of Production Function
- 7.9 Summary
- 7.10 Self - Assessment Questions
- 7.11 Further Readings

7.1 INTRODUCTION

Production process involves the transformation of inputs into output. The inputs could be land, labour, capital, entrepreneurship etc. and the output could be goods or services. In a production process managers take four types of decisions: (a) whether to produce or not, (b) how much output to produce, (c) what input combination to use, and (d) what type of technology to use. This Unit deals with the analysis of managers' decision rules concerning (c) and (d) above. The analysis of the other two decisions will be covered in Units 8 and Unit 9 of this block.

In this unit, we shall begin with a general discussion of the concept of production function. The analysis of this unit mainly focuses on the firms that produce a single product. Analysis on decisions related to multiproduct firms is also given briefly. The nature of production when there is only one variable input is taken up first. We then move on to the problem of finding optimum

combination of inputs for producing a particular level of output when there are two or more variable inputs. You will also learn about the production decisions in case of product mix of multiproduct firms. The unit concludes with various functional forms of production frequently used by economists and their empirical estimation.

7.2 PRODUCTION FUNCTION

Suppose we want to produce apples. We need land, seedlings, fertilizer, water, labour, and some machinery. These are called *inputs* or *factors of production*. The *output* is apples. In general a given output can be produced with different combinations of inputs. A *production function* is the functional relationship between inputs and output. It shows the maximum output which can be obtained for a given combination of inputs. It expresses the *technological relationship* between inputs and output of a product.

In general, we can represent the production function for a firm as:

$$Q = f(x_1, x_2, \dots, x_n)$$

Where Q is the maximum quantity of output, $x_1, x_2, \dots, x_n$ are the quantities of various inputs, and f stands for functional relationship between inputs and output. For the sake of clarity, let us restrict our attention to only one product produced using either one input or two inputs. If there are only two inputs, capital (K) and labour (L), we write the production function as:

$$Q = f(L, K)$$

This function defines the maximum rate of output (Q) obtainable for a given rate of capital and labour input. It may be noted here that outputs may be tangible like computers, television sets, etc., or it may be intangible like education, medical care, etc. Similarly, the inputs may be other than capital and labour. Also, the principles discussed in this unit apply to situations with more than two inputs as well.

Economic Efficiency and Technical Efficiency

We say that a firm is *technically efficient* when it obtains maximum level of output from any given combination of inputs. The production function incorporates the technically efficient method of production. A producer cannot decrease one input and at the same time maintain the output at the same level without increasing one or more inputs. When economists use production functions, they assume that the maximum output is obtained from any given combination of inputs. That is, they assume that production is technically efficient.

On the other hand, we say a firm is *economically efficient*, when it produces a given amount of output at the lowest possible cost for a combination of inputs

provided that the prices of inputs are given. Therefore, when only input combinations are given, we deal with the problem of technical efficiency; that is, how to produce maximum output. On the other hand, when input prices are also given in addition to the combination of inputs, we deal with the problem of economic efficiency; that is, how to produce a given amount of output at the lowest possible cost.

One has to be careful while interpreting whether a production process is efficient or inefficient. Certainly a production process can be called efficient if another process produces the same level of output using one or more inputs, other things remaining constant. However, if a production process uses less of some inputs and more of others, the economically efficient method of producing a given level of output depends on the prices of inputs. Even when two production processes are technically efficient, one process may be economically efficient under one set of input prices, while the other production process may be economically efficient at other input prices.

Let us take an example to differentiate between technical efficiency and economic efficiency. An ABC company is producing readymade garments using cotton fabric in a certain production process. It is found that 10 percent of fabric is wasted in that process. An engineer suggested that the wastage of fabric can be eliminated by modifying the present production process. To this suggestion, an economist reacted differently saying that if the cost of wasted fabric is less than that of modifying production process then it may not be economically efficient to modify the production process.

Short Run and Long Run

All inputs can be divided into two categories; i) fixed inputs and ii) variable inputs. A fixed input is one whose quantity cannot be varied during the time under consideration. The time period will vary depending on the circumstances. Although any input may be varied no matter how short the time interval, the cost involved in augmenting the amount of certain inputs is enormous; so as to make quick variation impractical. Such inputs are classified as fixed and include plant and equipment of the firm.

On the other hand, a variable input is one whose amount can be changed during the relevant period. For example, in the construction business the number of workers can be increased or decreased on short notice. Many 'builder' firms employ workers on a daily wage basis and frequent change in the number of workers is made depending upon the need. The amount of milk that goes in the production of butter can be altered quickly and easily and is thus classified as a variable input in the production process.

Whether or not an input is fixed or variable depends upon the time period involved. The longer the length of the time period under consideration, the more likely it is that the input will be variable and not fixed. Economists find it convenient to distinguish between the short run and the long run. The short

run is defined to be that period of time when some of the firm's inputs are fixed. Since it is most difficult to change plant and equipment among all inputs, the short run is generally accepted as the time interval over which the firm's plant and equipment remain fixed. In contrast, the long run is that period over which all the firms' inputs are variable. In other words, the firm has the flexibility to adjust or change its environment.

Production processes of firms generally permit a variation in the proportion in which inputs are used. In the long run, input proportions can be varied considerably. For example, at Maruti Udyog Limited, an automobile dye can be made on conventional machine tools with more labour and less expensive equipment, or it can be made on numerically controlled machine tools with less labour and more expensive equipment i.e. the amount of labour and amount of equipment used can be varied. Later in this unit, this aspect is considered in more detail. On the other hand, there are very few production processes in which inputs have to be combined in fixed proportions. Consider, Ranbaxy or Smith-Kline-Beecham or any other pharmaceutical firm. In order to produce a drug, the firm may have to use a fixed amount of aspirin per 10 gm of the drug. Even in this case a certain (although small) amount of variation in the proportion of aspirin may be permissible. If, on the other hand, no flexibility in the ratio of inputs is possible, the technology is described as fixed proportion type. We refer to this extreme case later in this unit, but as should be apparent, it is extremely rare in practice.

Activity 1

1. What is a production function? How does a long run production function differ from a short run production function?

.....

.....

.....

.....

.....

.....

2. When can we say that a firm is: (a) technically efficient, (b) economically efficient? Is it necessary that a technically efficient firm is also economically efficient?

.....

.....

.....

.....

.....

.....

7.3 PRODUCTION FUNCTION WITH ONE VARIABLE INPUT

Consider the simplest two input production process - where one input with a fixed quantity and the other input with its variable quantity. Suppose that the fixed input is the service of machine tools, the variable input is labour, and the output is a metal part. The production function in this case can be represented as:

$$Q = f(K, L)$$

Where Q is output of metal parts, K is service of five machine tools (fixed input), and L is labour (variable input). The variable input can be combined with the fixed input to produce different levels of output.

Total, Average, and Marginal Products

The production function given above shows us the maximum total *product* (TP) that can be obtained using different combinations of quantities of inputs. Suppose the metal parts company decides to know the output level for different input levels of labour using fixed five machine tools. Table 7.1 explains the total output for different levels of variable input. In this example, the TP rises with increase in labour up to a point (six workers), becomes constant between sixth and seventh workers, and then declines.

**Table 7.1: Total, Average and Marginal Products of labour
(with fixed capital at five machine tools)**

Number of workers (L)	Total output (TP) (thousands per year) (Q)	Marginal product ($MP_L = \Delta Q / \Delta L$)	Average product ($AP_L = Q / L$)
0	0	—	—
1	10	10	10
2	28	18	14
3	54	26	18
4	76	22	19
5	90	14	18
6	96	6	16
7	96	0	13.5
8	92	-4	11.5

Two other important concepts are the *average product* (AP) and the *marginal product* (MP) of an input. The AP of an input is the TP divided by the amount of input used to produce this amount of output. Thus AP is the output-input ratio for each level of variable input usage. The MP of an input is the addition to TP resulting from the addition of one unit of input, when the amounts of other inputs are constant. In our example of machine parts

production process, the AP of labour is the TP divided by the number of workers.

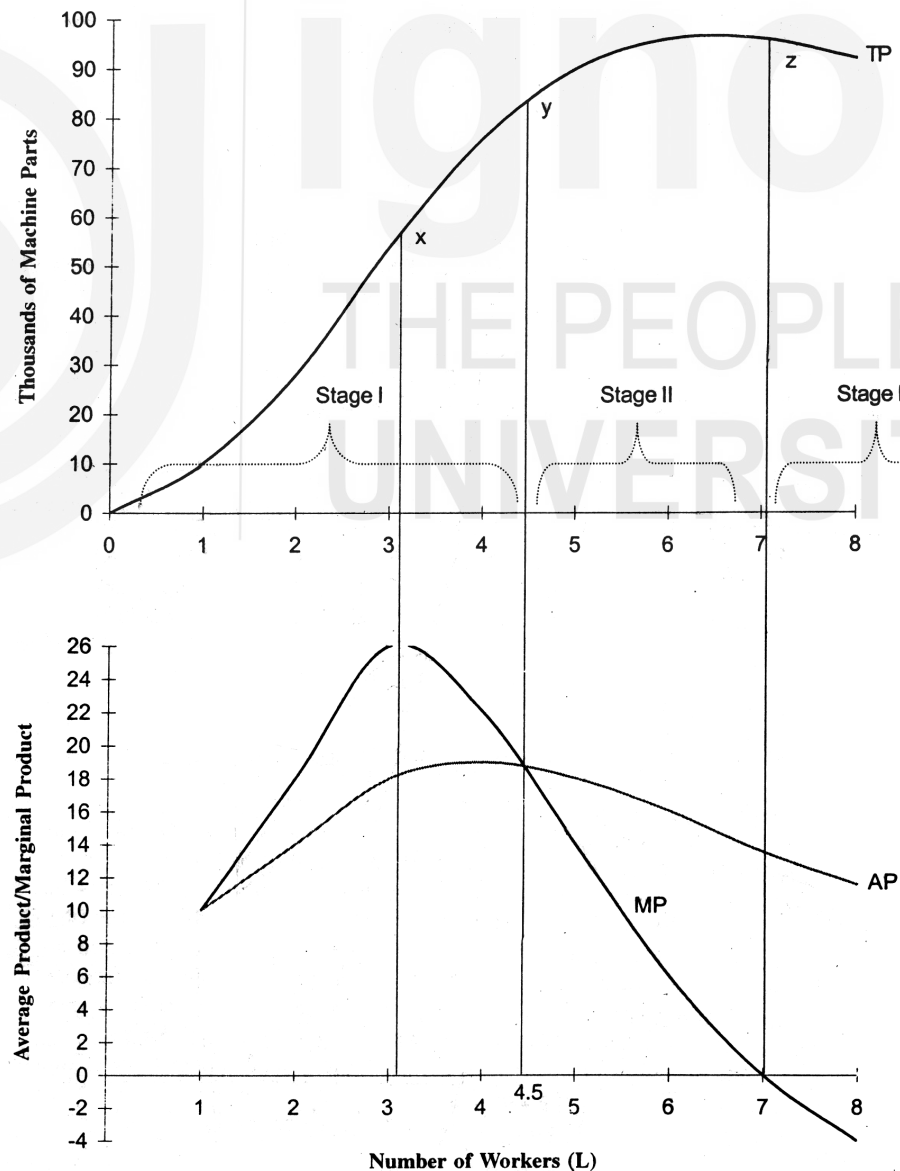
$$AP_L = Q/L$$

As shown in Table 7.1, the AP_L first rises, reaches maximum at 19, and then declines thereafter. Similarly, the MP of labour is the additional output attributable to using one additional worker with use of other input (service of five machine tools) fixed.

$$MP_L = \Delta Q / \Delta L$$

Where Δ means 'the change in'. For example, from Table 7.1 for MP_4 (marginal product of 4th worker) $\Delta Q = 76 - 54 = 22$ and $\Delta L = 4 - 3 = 1$. Therefore, $MP_4 = (22/1) = 22$. Note that although the MP first increases with addition of workers, it declines later and for the addition of 8th worker it becomes negative (-4).

Figure 7.1: Relationship between TP, MP, and AP curves and the three stages of production



The graphical presentation of total, average, and marginal products for our example of machine parts production process is shown in Figure 7.1.

Relationship between TP, MP and AP Curves

Examine Table 7.1 and its graphical presentation in Figure 7.1. We can establish the following relationship between TP, MP, and AP curves.

- 1a) If $MP > 0$, TP will be rising as L increases. The TP curve begins at the origin, increases at an increasing rate over the range 0 to 3, and then increases at a decreasing rate. The MP reaches a maximum at 3, which corresponds to an inflection point (x) on the TP curve. At the inflection point, the TP curve changes from increasing at an increasing rate to increasing at decreasing rate.
- b) If $MP = 0$, TP will be constant as L increases. The TP is constant between workers 6 and 7.
- c) If $MP < 0$, TP will be declining as L increases. The TP declines beyond 7. Also, the TP curve reaches a maximum when $MP = 0$ and then starts declining when $MP < 0$.
2. MP intersects AP ($MP = AP$) at the maximum point on the AP curve. This occurs at labour input rate 4.5. Also, observe that whenever $MP > AP$, the AP is rising (upto number of workers 4.5) — it makes no difference whether MP is rising or falling. When $MP < AP$ (from number of workers 4.5), the AP is falling. Therefore, the intersection must occur at the maximum point of AP. It is important to understand why. The key is that AP increases as long as the MP is greater than AP. And AP decreases as long as MP is less than AP. Since AP is positively or negatively sloped depending on whether MP is above or below AP, it follows that $MP=AP$ at the highest point on the AP curve.

This relationship between MP and AP is not unique to economics. Consider a cricket batsman, say Sachin Tendulkar, who is averaging 50 runs in 10 innings. In his next innings he scores a 100. His marginal score is 100 and his average will now be above 50. More precisely, it is 54 i.e. $(50 * 10 + 100) / (10+1) = 600/11$. This means when the marginal score is above the average, the average must increase. In case he had scored zero, his marginal score would be below the average, and his average would fall to 45.5 i.e. $500/11$ is 45.45. Only if he had scored 50 would the average remain constant, and the marginal score would be equal to the average.

The Law of Diminishing Marginal Returns

The slope of the MP curve in Figure 7.1 illustrates an important principle, *the law of diminishing marginal returns*. As the number of units of the variable input increases, the other inputs held constant (fixed), there exists a point beyond which the MP of the variable input declines. Table 7.1 illustrates this law. Observe that MP was increasing up to the addition of 4th worker (input);

beyond this the MP decreases. What this law says is that MP may rise or stay constant for some time, but as we keep increasing the units of variable input, MP should start falling. It may keep falling and turn negative, or may stay positive all the time. Consider another example for clarity. Single application of fertilizers may increase the output by 50%, a second application by another 30% and the third by 20% and so on. However, if you were to apply fertilizer five to six times in a year, the output may drop to zero.

Three things should be noted concerning the law of diminishing marginal returns.

1. This law is an empirical generalization, not a deduction from physical or biological laws.
2. It is assumed that technology remains fixed. The law of diminishing marginal returns cannot predict the effect of an additional unit of input when technology is allowed to change.
3. It is assumed that there is at least one input whose quantity is being held constant (fixed). In other words, the law of diminishing marginal returns does not apply to cases where all inputs are variable.

Stages of Production

Based on the behaviour of MP and AP, economists have classified production into three stages:

Stage 1: $MP > 0$, AP rising. Thus, $MP > AP$.

Stage 2: $MP > 0$, but AP is falling. $MP < AP$ but TP is increasing (because $MP > 0$).

Stage 3: $MP < 0$. In this case TP is falling.

These results are illustrated in Figure 7.1. No profit-maximizing producer would produce in stages I or III. In stage I, by adding one more unit of labour, the producer can increase the AP of all units. Thus, it would be unwise on the part of the producer to stop the production in this stage. As for stage III, it does not pay the producer to be in this region because by reducing the labour input the total output can be increased and the cost of a unit of labour can be saved.

Thus, the economically meaningful range is given by stage II. In Figure 7.1 at the point of inflection (x), we saw earlier that MP is maximized. At point y, since AP is maximized, we have $AP = MP$. At point z, TP reaches a maximum. Thus, $MP = 0$ at this point. If the variable input is free then the optimum level of output is at point z where TP is maximized. However, in practice no input will be freely available. The producer has to pay a price for it. Suppose the producer pays Rs. 200 per worker per day and the price of a

unit of output (say one apple) is Rs. 10. In this case the producer will keep on hiring additional workers as long as

$(\text{Price of a unit of output}) * (\text{marginal product of labour}) > (\text{price of a unit of labour})$

That is, marginal revenue of product (MRP) of labour $> P_L$

On a similar analogy,

$(\text{price of a unit of output}) * (\text{marginal product of capital}) > (\text{price of a unit of capital})$

That is, marginal revenue of product (MRP) of capital $> P_K$

The left side denotes the increase in revenue and the right side denotes the increase in the cost of adding one more unit of labour. As long as the increment to revenues exceeds the increment to costs, the profit of the producer will increase. As we increase the units of labour, we see that MP diminishes. We assume that the prices of inputs and output do not change. In this case, as MP declines, revenues will start falling, and a point will come when the increase in revenue equals the increase in cost. At this point the producer will stop adding more units of input. With further addition, since MP declines, the additional revenues would be less than the additional costs, and the profit of the producer would decline.

Thus, profit maximization implies that a producer with no control over prices will increase the use of an input until—

Value of marginal product (MP) = Price of a unit of variable input

Activity 2

1. Fill in the blanks of the following Table

Capital	Labour	TP	AP_L	MP_L
1	0	0		
1	1	2	2	
1	2	5		3
1	3		3	4
1	4	12	3	
1	5	14		
1	6		$2\frac{1}{2}$	1
1		$15\frac{3}{4}$	$2\frac{1}{4}$	
1	8	11		
1	9		1	-2

2. State clearly the relation between AP_L and MP_L .

.....

.....

.....

.....

.....

.....

3. Why is the marginal product of labour likely to increase and then decline in the short-run?

.....

.....

.....

.....

.....

.....

4. Faced with constantly changing conditions, why would a firm ever keep any factors fixed? What determines whether a factor is fixed or variable?

.....

.....

.....

.....

.....

.....

5. Suppose a chair manufacturer is producing in the short run where equipment is fixed. The manufacturer knows that as the number of labourers used in the production process increases from 1 to 7, the number of chairs produced changes as follows: 10, 17, 22, 25, 26, 25, and 23.

- a) Calculate the marginal and average product of labour for this production function.
- b) Does this production function exhibit increasing returns to labour or decreasing returns to labour or both? Explain.

- c) Explain intuitively what might cause the marginal product of labour to become negative?

.....

.....

.....

.....

.....

.....

6. Why a profit-maximizing producer would produce in stage-II and not in stage-I or III? Explain.

.....

.....

.....

.....

.....

.....

7.4 PRODUCTION FUNCTION WITH TWO VARIABLE INPUTS

Now we turn to the case of production where two inputs (say capital and labour) are variable. Although, we restrict our analysis to two variable inputs, all of the results hold for more than two also. We are restricting our analysis to two variable inputs because it simply allows us the scope for graphical analysis. When analyzing production with more than one variable input, we cannot simply use sets of AP and MP curves like those discussed in section 7.3, because these curves were derived holding the use of all other inputs fixed and letting the use of only one input vary. If we change the level of fixed input, the TP, AP and MP curves would shift. In the case of two variable inputs, changing the use of one input would cause a shift in the MP and AP curves of the other input. For example, an increase in capital would probably result in an increase in the MP of labour over a wide range of labour use.

Production Isoquants

In Greek the word 'iso' means 'equal' or 'same'. A production *isoquant* (*equal output curve*) is the locus of all those combinations of two inputs which yields a given level of output. With two variable inputs, capital and labour, the isoquant gives the different combinations of capital and labour, that produces the same level of output. For example, 5 units of output can be produced using either 15 units of capital (K) or 2 units of labour (L) or $K=10$ and $L=3$ or $K=5$ and $L=5$ or

$K=3$ and $L=7$. These four combinations of capital and labour are four points on the isoquant associated with 5 units of output as shown in Figure 7.2. And if we assume that capital and labour are continuously divisible, there would be many more combinations on this isoquant.

Now let us assume that capital, labour, and output are continuously divisible in order to set forth the typically assumed characteristics of isoquants. Figure 7.3 illustrates three such isoquants. Isoquant I shows all the combinations of capital and labour that will produce 10 units of output. According to this isoquant, it is possible to obtain this output if K_0 units of capital and L_0 units of

Figure 7.2: Production Isoquant: This isoquant shows various combinations of capital and labour inputs that can produce 5 units of output.

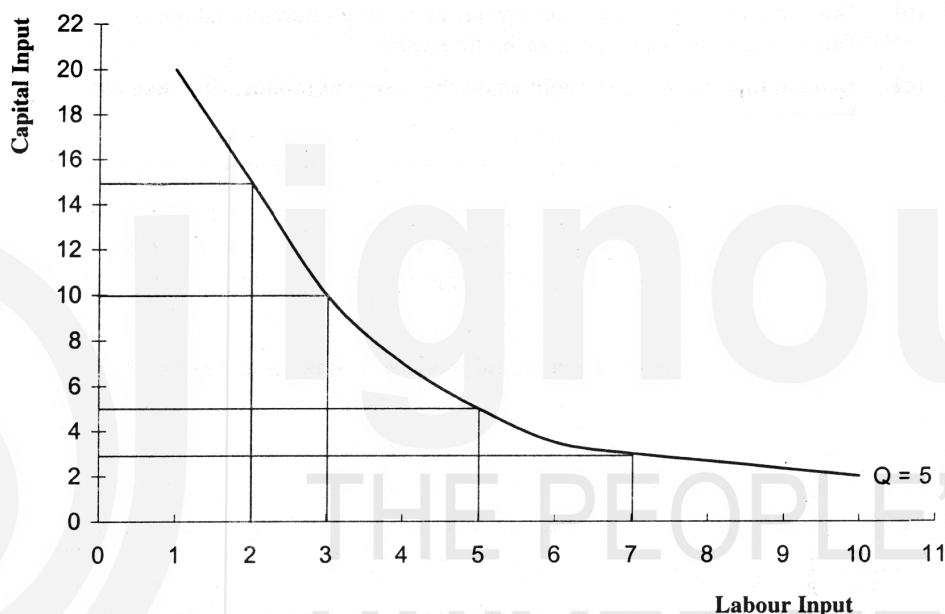
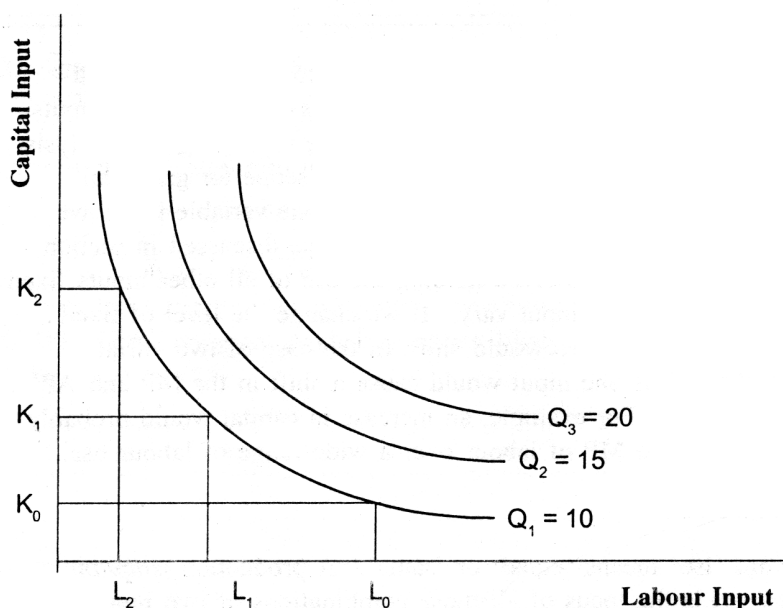


Figure 7.3: Isoquant Map: These isoquants shows various combinations of capital and labour inputs that can produce 10, 15, and 20 units of output.



labour inputs are used. Alternately, this output can also be obtained if K_1 units of capital and L_1 units of labour inputs or K_2 units of capital and L_2 units of labour are used. Similarly, isoquant II shows the various combinations of capital and labour that can be used to produce 15 units of output. Isoquant III shows all combinations that can produce 20 units of output. Each capital-labour combination can be on only one isoquant. That is, isoquants cannot intersect. These isoquants are only three of an infinite number of isoquants that could be drawn. A group of isoquants is called an isoquant map. In an isoquant map, all isoquants lying above and to the right of a given isoquant indicate higher levels of output. Thus, in Figure 7.3 isoquant II indicates a higher level of output than isoquant I, and isoquant III indicates a higher level of output than isoquant II.

In general, isoquants are determined in the following way. First, a rate of output, say Q_0 , is specified. Hence the production function can be written as

$$Q_0 = f(K, L)$$

Those combinations of K and L that satisfy this equation define the isoquant for output rate Q_0 .

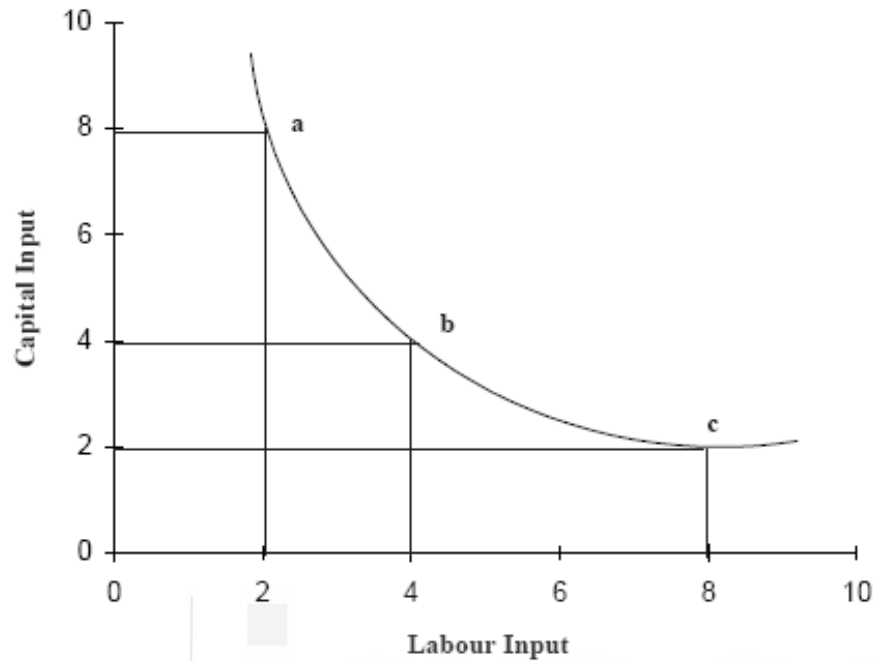
Marginal Rate of Technical Substitution

As we have seen above, generally there are a number of ways (combinations of inputs) that a particular output can be produced. The rate, at which one input can be substituted for another input, if output remains constant, is called the *marginal rate of technical substitution* (MRTS). It is defined in case of two inputs, capital and labour, as the amount of capital that can be replaced by an extra unit of labour, without affecting total output.

$$MRTS_{L \text{ for } K} = \left| \frac{\Delta K}{\Delta L} \right|$$

It is customary to define the MRTS as a positive number, since $\Delta K / \Delta L$, the slope of the isoquant, is negative. Over the relevant range of production the MRTS diminishes. That is, more and more labour is substituted for capital while holding output constant, the absolute value of $\Delta K / \Delta L$ decreases. For example, let us assume that 10 pairs of shoes can be produced using either 8 units of capital and 2 units of labour or 4 units each of capital and of labour or 2 units of capital and 8 units of labour. From Figure 7.4 the MRTS of labour for capital between points a and b is equal to $\Delta K / \Delta L = (4-8) / (4-2) = -4/2 = -2$ or $|2|$. Between points b and c, the MRTS is equal to $-2/4 = -1/2$ or $|1/2|$. The MRTS has decreased because capital and labour are not perfect substitutes for each other. Therefore, as more of labour is added, less of capital can be used (in exchange for another unit of labour) while keeping the output level constant.

Figure 7.4: Marginal Rate of Technical Substitution



There is a simple relationship between MRTS of labour for capital and the marginal product MP_K and MP_L of capital and labour respectively. Since along an isoquant, the level of output remains the same, if ΔL units of labour are substituted for ΔK units of capital, the increase in output due to ΔL units of labour (namely, $\Delta L * MP_L$) should match the decrease in output due to a decrease of ΔK units of capital (namely, $\Delta K * MP_K$). In other words, along an isoquant,

$$\Delta L * MP_L = \Delta K * MP_K$$

Which is equal to

$$\left| \frac{\Delta K}{\Delta L} \right| = \frac{MP_L}{MP_K}$$

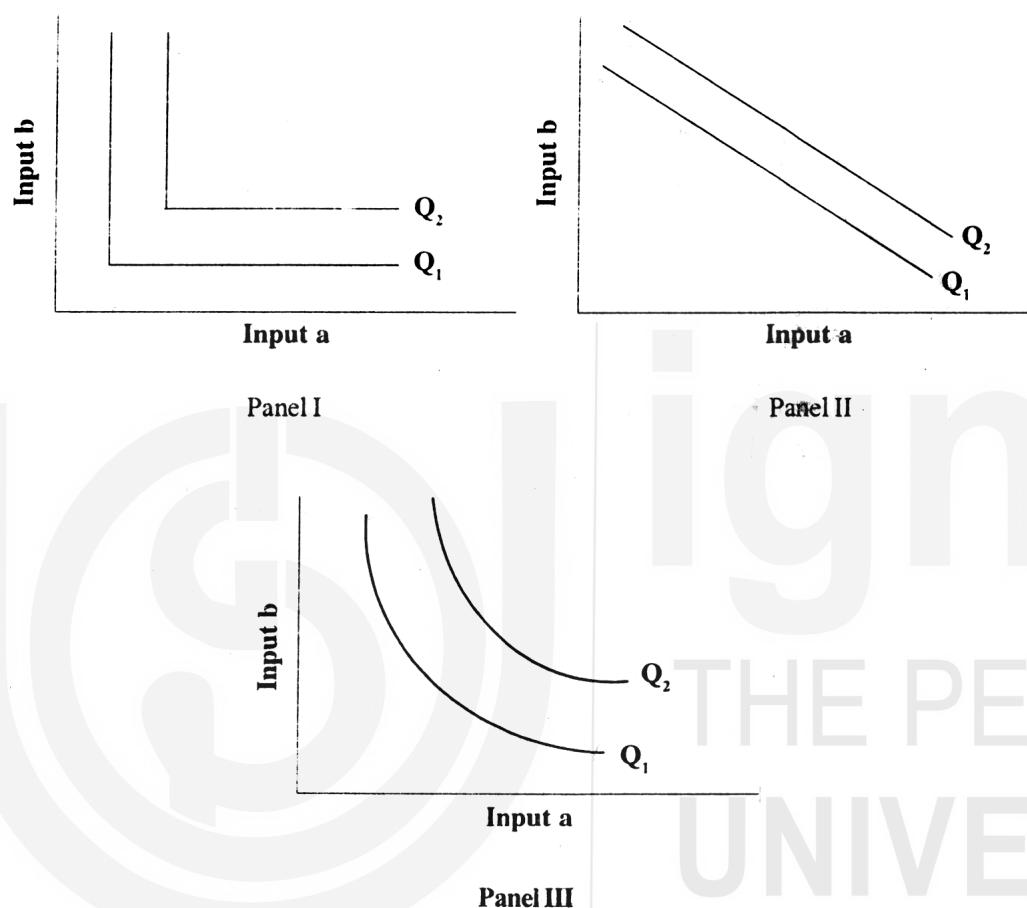
However, as we have seen earlier $\Delta K/\Delta L$ is equal to $MRTS_L$ for K, and hence, we get the following expression for MRTS of L for K as the ratio of the corresponding marginal products.

$$MRTS_{L \text{ For } K} = \frac{MP_L}{MP_K}$$

There are vast differences among inputs in how readily they can be substituted for one another. For example, in some extreme production process, one input can perfectly be substituted for another; whereas in some other extreme production process no substitution is possible. On the other hand, in most of the production processes what we see is imperfect substitution of inputs. These three general shapes that an isoquant might have are shown in Figure 7.5. In panel I, the isoquants are right angles implying that the two inputs a and b must be used in fixed proportion and

they are not at all substitutable. For instance, there is no substitution possible between the tyres and a battery in an automobile production process. The MRTS in all such cases would, therefore, be zero. The other extreme case would be where the inputs a and b are perfect substitutes as shown in panel II. The isoquants in this category will be a straight line with constant slope or MRTS. A good example of this type

Figure 7.5: Three General Types of Shapes of Isoquants



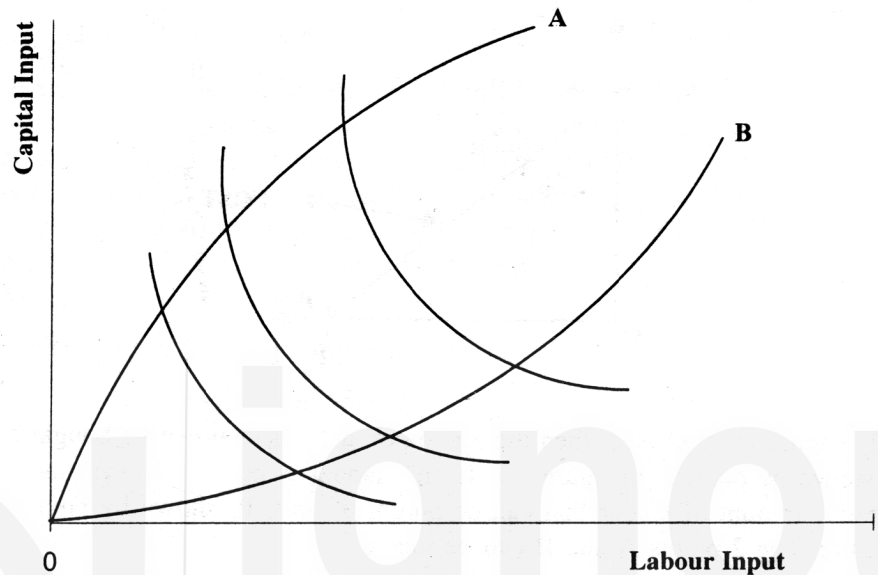
would be natural gas and fuel oil, which are close substitutes in energy production. The most common situation is presented in panel III. The inputs are imperfect substitutes in this case and the rate at which input a can be given up in return for one more unit of input b keeping the output constant diminishes as the amount of input b increases.

The Economic Region of Production

Isoquants may also have positively sloped segments, or bend back upon themselves, as shown in Figure 7.6. Above OA and below OB , the slope of the isoquants is positive, which implies that increase in both capital and labour are required to maintain a certain output rate. If this is the case, the MP of one or other input must be negative. Above OA , the MP of capital is negative. Thus output will increase if less capital is used, while the amount of labour is held constant. Below OB , the MP of labour is negative.

Thus, output will increase if less labour is used, while the amount of capital is held constant. The lines OA and OB are called *ridge lines*. And the region bounded by these ridge lines is called *economic region* of production. This means the region of production beyond the ridge lines is *economically inefficient*.

Figure 7.6: Economic Region of Production



7.5 THE OPTIMAL COMBINATION OF INPUTS

In the above section you have learned that any desired level of output can be produced using a number of different combinations of inputs. As said earlier in the introduction of this unit one of the decision problems that concerns a production process manager is, which input combination to use. That is, what is the *optimal input combination*? While all the input combinations are technically efficient, the final decision to employ a particular input combination is purely an economic decision and rests on cost (expenditure). Thus, the production manager can make either of the following two input choice decisions:

1. Choose the input combination that yields the maximum level of output with a given level of expenditure.
2. Choose the input combination that leads to the lowest cost of producing a given level of output.

Thus, the decision is to minimize cost subject to an output constraint or maximize the output subject to a cost constraint. We will now discuss these two fundamental principles. Before doing this we will introduce the concept *isocost*, which shows all combinations of inputs that can be used for a given cost.

Isocost Lines

Recall that a universally accepted objective of any firm is to maximize profit. If the firm maximizes profit, it will necessarily minimize cost for producing a

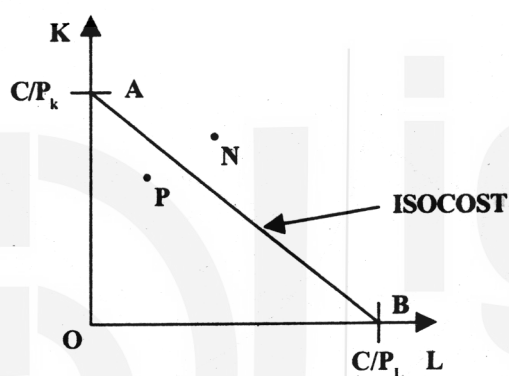
given level of output or maximize output for a given level of cost. Suppose there are 2 inputs: capital (K) and labour (L) that are variable in the relevant time period. What combination of (K,L) should the firm choose in order to maximize output for a given level of cost?

If there are 2 inputs, K,L, then given the price of capital (P_K) and the price of labour (P_L), it is possible to determine the alternative combinations of (K,L) that can be purchased for a given level of expenditure. Suppose C is total expenditure, then

$$C = P_L * L + P_K * K$$

This linear function can be plotted on a graph.

Figure 7.7 : Isocost line



If only capital is purchased, then the maximum amount that can be bought is C/P_K shown by point A in figure 7.7. If only labour is purchased, then the maximum amount of labour that can be purchased is C/P_L shown by point B in the figure. The 2 points A and B can be joined by a straight line. This straight line is called the **isocost line** or equal cost line. It shows the alternative combinations of (K,L) that can be purchased for the given expenditure level C. Any point to the right and above the isocost is not attainable as it involves a level of expenditure greater than C and any point to the left and below the isocost such as P is attainable, although it implies the firm is spending less than C. You should verify that the slope of the isocost is¹

$$-\frac{\Delta K}{\Delta L} = \frac{C}{P_K} * \frac{C}{P_L} = \frac{P_L}{P_K}$$

EXAMPLE

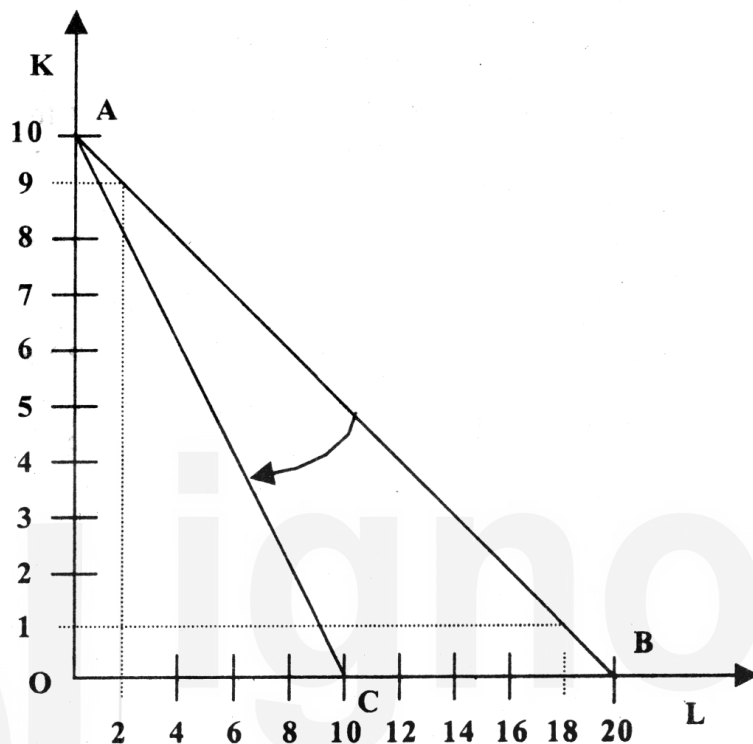
Consider the following data:

$$P_L = 10, \quad P_K = 20 \quad \text{Total Expenditure} = 200$$

¹ The negative sign is due to the fact that the slope of the isocost is negative.

Let us first plot the various combinations of K and L that are possible. We consider only the case when the firm spends the entire budget of 200. The alternative combinations are shown in the figure (7.8).

Figure 7.8 : Shifting of Iscost



The slope of this isocost is $-\frac{1}{2}$. What will happen if labour becomes more expensive say P_L increases to 20? Obviously with the same budget the firm can now purchase lesser units of labour. The isocost still meets the Y-axis at point A (because the price of capital is unchanged), but shifts inwards in the direction of the arrow to meet the X-axis at point C. The slope therefore changes to -1 . You should work out the effect on the isocost curve on the following:

- i) decrease in the price of labour
- ii) increase in the price of capital
- iii) decrease in the price of capital
- iv) increase in the firms budget with no change in the price of labour and capital

[Hint: The slope of the isocost will not change in this case]

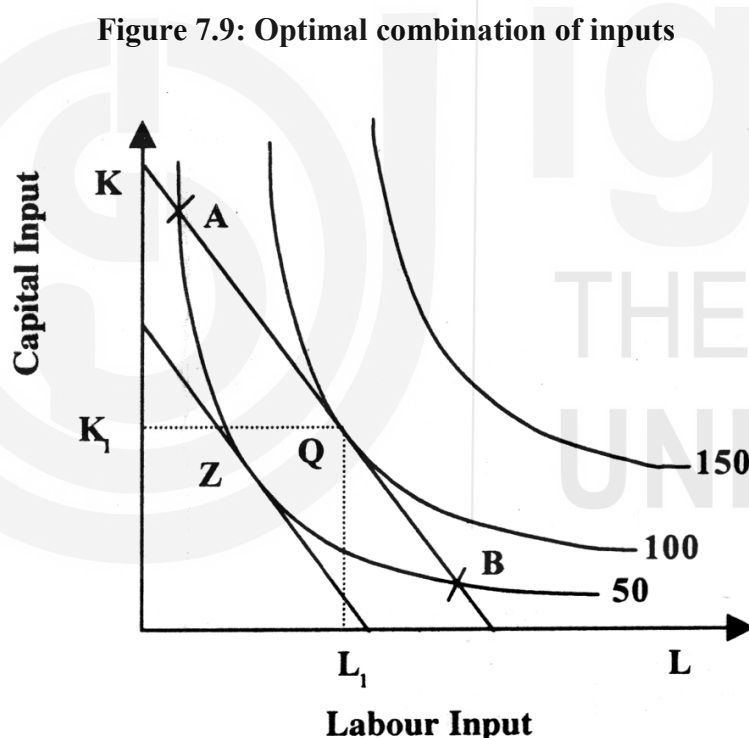
Optimal Combination of Inputs: The Long Run

When both capital and labour are variable, determining the optimal input rates of capital and labour requires the technical information from the production function i.e. the isoquants be combined with market data on input prices i.e. the isocost function. If we superimpose the relevant isocost curve

on the firm's isoquant map, we can readily determine graphically as to which combination of inputs maximize the output for a given level of expenditure.

Consider the problem of minimizing the cost of a given rate of output. Specifically if the firm wants to produce 50 units of output at minimum cost. Two production isoquants have been drawn in Figure 7.9. Three possible combinations (amongst a number of more combinations) are indicated by points A, Z and B in Figure 7.9. Obviously, the firm should pick the point on the lower isocost, i.e. point Z. In fact, Z is the minimum cost combination of capital and labour. A to Z the isocost is tangent to the 50 unit isoquant.

Alternatively, consider the problem of maximizing output subject to a given cost amount. You should satisfy yourself that among all possible output levels, the maximum amount will be represented by the isoquant that is tangent to the relevant isocost line. Suppose the budget of the firm increases to the amount shown by the higher of the two isocost lines in Figure 7.9, point Q or 100 units of output is the maximum attainable given the new cost constraint in Figure 7.9.



Regardless of the production objective, efficient production requires that the isoquant be tangent to the isocost function. If the problem is to maximize output, subject to a cost constraint or to minimize cost for a given level of output, the same efficiency condition holds true in both situations. Intuitively, if it is possible to substitute one input for another to keep output constant while reducing total cost, the firm is not using the least cost combination of inputs. In such a situation, the firm should substitute one input for another.

For example, if an extra rupee spent on capital generates more output than an extra rupee spent on labour, then more capital and less labour should be employed. At point Q in Figure 7.9, the marginal product of capital per rupee

spent on capital is equal to the marginal product of labour per rupee spent on labour. Mathematically this can be shown as

$$\frac{MP_L}{P_L} = \frac{MP_K}{P_K} \dots\dots\dots 1$$

Or equivalently,

$$\frac{MP_L}{MP_K} = \frac{P_L}{P_K} \dots\dots\dots 2$$

Whenever the 2 sides of the above equation are not equal, there are Possibilities that input substitutions will reduce costs. Let us work with numbers. Suppose $P_L = 10, P_K = 20$,

$MP_L = 50$ and $MP_K = 40$. Thus, we have

$$\frac{50}{10} > \frac{40}{20}$$

This cannot be an efficient input combination, because the firm is getting more output per rupee spent on labour than on capital. If one unit of capital is sold to obtain 2 units of labour ($P_K = 20, P_L = 10$), net increase in output will be 160. Thus the substitution of labour for capital would result in a net increase in output at no additional cost. The inefficient combination corresponds to a point such as A in Figure 7.9. At that point too much capital is employed. The firm, in order to maximize profits will move down the isocost line by substituting labour for capital until it reaches point Q. Conversely, at a point such as B in figure 7.9 the reverse is true—there is too much labour and the inequality

$$\frac{MP_L}{P_L} < \frac{MP_K}{P_K} \text{ will hold}$$

¹Recall that $\frac{MP_L}{MP_K}$ is the slope of the isoquant and it is also the MRTS while $\frac{P_L}{P_K}$ is the slope of the isocost line. Since for optimum, the isocost must be tangent to the isoquant, the result follows. Many text books denote P_L which is the price of labour as w or the wage rate and P_K which is the price of capital as r or the rental. The equilibrium condition can thus also be written as

$$\frac{MP_L}{MP_K} = \frac{w}{r}$$

² Since the $MP_L = 50$, 2 units of labour produce 100 units, while reducing capital by 1 unit decreases output by 40 units ($MP_K = 40$). Therefore, net increase is 60 units. This, of course, assumes that MP_L and MP_K remain constant in the relevant range. We know that as more labour is employed in place of capital, MP_L will decline and MP_K will increase (this follows from the law of diminishing returns) and thus equation (2) will be satisfied.

This means that the firm generates more output per rupee spent on capital than from rupees spent on labour. Thus a profit maximizing firm should substitute capital for labour.

Suppose the firm was operating at point B in Figure 7.9. If the problem is to minimize cost for a given level of output (B is on the isoquant that corresponds to 50 units of output), the firm should move from B to Z along the 50-unit isoquant thereby reducing cost, while maintaining output at 50. Alternatively, if the firm wants to maximize output for given cost, it should move from B to Q, where the isocost is tangent to the 100-unit isoquant. In this case output will increase from 50 to 100 at no additional cost. Thus both the following decisions

- (a) the input combination that yields the maximum level of output with a given level of expenditure, and
- (b) the input combination that leads to the lowest cost of producing a given level of output are satisfied at point Q in Figure 7.9.

You should be satisfied that this is indeed the case.

The isocost-isoquant framework described above lends itself to various applications. It demonstrates, simply and elegantly, when relative prices of inputs change, managers will respond by substituting the input that has become relatively less expensive for the input that has become relatively more expensive. On average, we know that compared to developed countries like the US, UK, Japan and Germany, labour in India is less expensive. It is not surprising therefore to find production techniques that on average, use more labour per unit of capital in India than in the developed world. For example, in construction activity you see around you in your city, inexpensive workers do the job that in developed countries is performed by machines.

One application of the isocost-isoquant framework frequently cited is the response of industry to the rapidly rising prices of energy products in the 1970s. (Remember the oil price shock of 1973 and again of 1979). Most prices of petrol and petroleum products increased across the world, and as our analysis suggests, firms responded by conserving energy by substituting other inputs for energy.

Activity 3

1. Draw an isoquant map using the information available in the following Table.

Production and Cost Analysis

Isoquant-I		Isoquant-II		Isoquant-III	
L	K	L	K	L	K
2	11	4	13	6	15
1	8	3	10	5	12
2	5	4	7	6	9
3	3	5	5	7	7
4	2.3	6	4.2	8	6.2
5	1.8	7	3.5	9	5.5
6	1.6	8	3.2	10	5.3
7	1.8	9	3.5	11	5.5

Capital Input

Labour Input

- 1a) Which one of the isoquants provides you with highest level of output and why?

.....

.....

.....

.....

.....

.....

.....

.....

- 1b) Take any one of the isoquants and compute $MRTS_{LK}$. What do you observe about computed MRTS? Explain the observed trend.

Isoquant.....		
L	K	$MRTS_{LK}$

Production Function

2. The marginal product of labour in the production of computer chips is 50 chips per hour. The marginal rate of technical substitution of hours of labour for hours of machine-capital is $\frac{1}{4}$. What is the marginal product of capital?

.....

.....

.....

.....

.....

3. What would the isoquants look like if all inputs were nearly perfect substitutes in a production process? What if there was near-zero substitutability between inputs?

.....

.....

.....

.....

.....

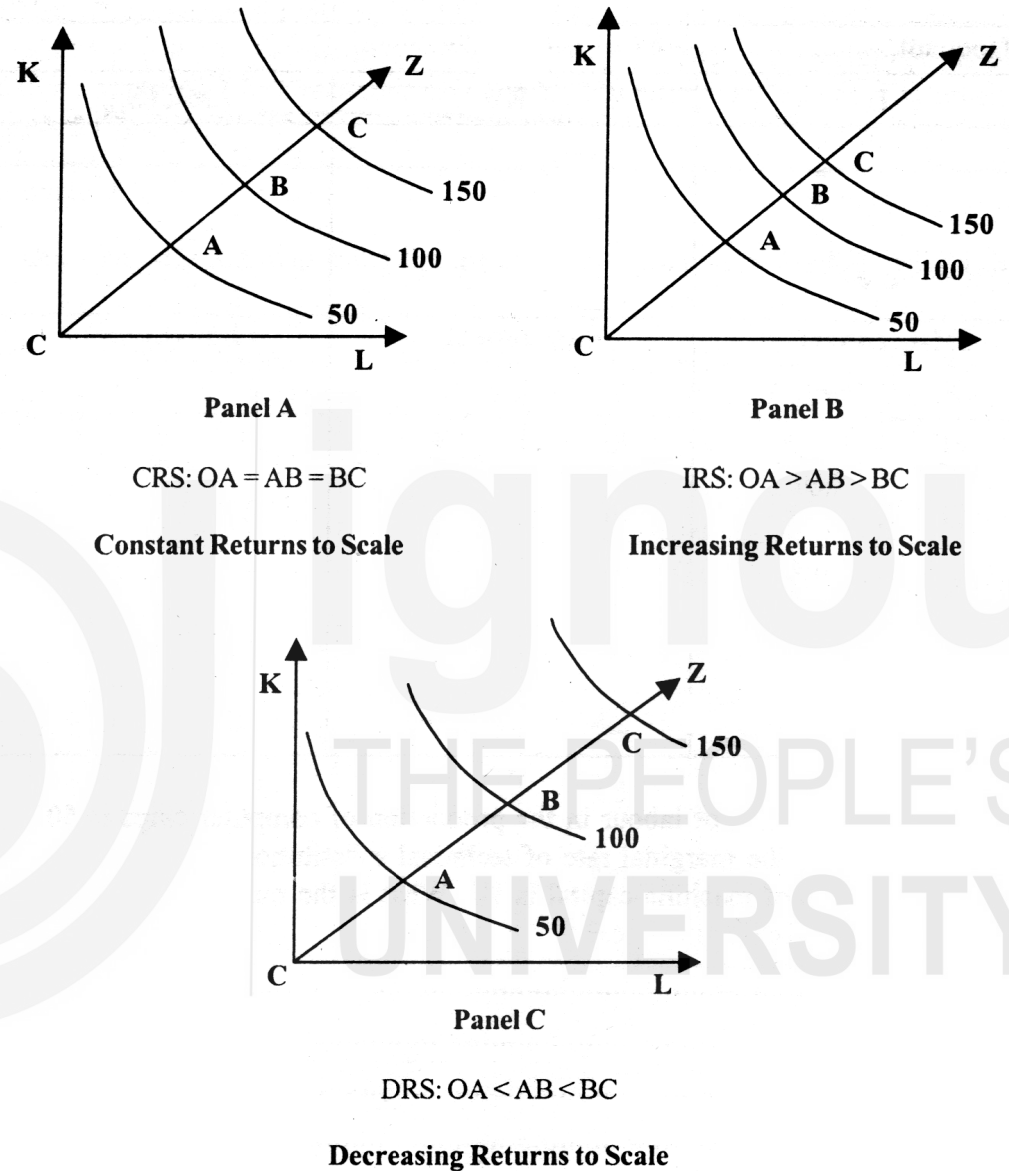
.....

7.6 RETURNS TO SCALE

Another important attribute of production function is how output responds in the long run to changes in the scale of the firm i.e. when all inputs are increased in the same proportion (by say 10%), how does output change. Clearly, there are 3 possibilities. If output increases by more than an increase in inputs (i.e. by more than 10%), then the situation is one of **increasing returns to scale (IRS)**. If output increases by less than the increase in inputs,

then it is a case of **decreasing returns to scale (DRS)**. Lastly, output may increase by exactly the same proportion as inputs. For example a doubling of inputs may lead to a doubling of B output. This is a case of **constant returns to scale (CRS)**.

Figure 7.10: Isoquants Showing Returns to Scale



In unit 8 we will examine returns to scale from the point of view of cost and also advance reasons for increasing and decreasing returns to scale. For the moment consider the following example. A box with dimensions $4 \times 4 \times 4$ has a capacity of 64 times a box with dimensions $1 \times 1 \times 1$, even though the former uses only 16 times more wood than the smaller box³.

³ Volume of box : $4 \times 4 \times 4 = 64$ Area of box: $4 \times 4 \times 6 = 96$
 $1 \times 1 \times 1 = 1$ $1 \times 1 \times 6 = 6$

Panel A shows constant returns to scale. Three isoquants with output levels 50, 100 and 150 are drawn. In the figure, successive isoquants are equidistant from one another along the ray CZ. Panel B shows increasing returns to scale, where the distance between 2 isoquants becomes less and less i.e. in order to double output from 50 to 100, input increase is less than double. The explanation for panel C, which exhibits decreasing returns to scale, is analogous.

There is no universal answer to which industries will show what kind of returns to scale. Some industries like public utilities (Telecom and Electricity generation) show increasing returns over large ranges of output, whereas other industries exhibit constant or even decreasing returns to scale over the relevant output range. Therefore, whether an industry has constant, increasing or decreasing returns to scale is largely an empirical issue.

7.7 FUNCTIONAL FORMS OF PRODUCTION FUNCTION

The production function can be estimated by regression techniques (refer to MMPC-005, course on “Quantitative Analysis for Managerial Applications” to know about regression techniques) using historical data (either time-series data, or cross-section data, or engineering data). For this, one of the first tasks is to select a functional form, that is, the specific relationship among the relevant economic variables. We know that the general form of production function is,

$$Q = f(K, L)$$

Where, Q = output, K = capital and L = labour.

Although, a variety of functional forms have been used to describe production relationships, only the Cobb-Douglas production function is discussed here. The general form of Cobb-Douglas function is expressed as:

$$Q = AK^\alpha L^\beta$$

where A, α , and β are the constants that, when estimated, describe the quantitative relationship between the inputs (K and L) and output (Q).

The marginal products of capital and labour and the rates of the capital and labour inputs are functions of the constants A, α , and β and. That is,

$$MP_K = \frac{dQ}{dK} = \alpha AK^{\alpha-1} L^\beta$$

$$MP_L = \frac{dQ}{dL} = \beta AK^\alpha L^{\beta-1}$$

The sum of the constants $(\alpha + \beta)$ can be used to determine returns to scale. That is,

$(\alpha + \beta) > 1 \Rightarrow$ increasing returns to scale,

$(\alpha + \beta) = 1 \Rightarrow$ constant returns to scale, and

$(\alpha + \beta) < 1 \Rightarrow$ decreasing returns to scale.

Having numerical estimates for the constants of the production function provides significant information about the production system under study. The marginal products for each input and returns to scale can all be determined from the estimated function.

The Cobb-Douglas function does not lend itself directly to estimation by the regression methods because it is a nonlinear relationship. Technically, an equation must be a linear function of the parameters in order to use the ordinary least-squares regression method of estimation. However, a linear equation can be derived by taking the logarithm of each term. That is,

$$\log Q = \log A + \alpha \log K + \beta \log L$$

A linear relationship can be seen by setting,

$$Y = \log Q, \quad A^* = \log A, \quad X_1 = \log K, \quad X_2 = \log L \text{ and}$$

Re writing the function as

$$Y = A^* + \alpha X_1 + \beta X_2$$

This function can be estimated directly by the least-squares regression technique and the estimated parameters used to determine all the important production relationships. Then the antilogarithm of both sides can be taken, which transforms the estimated function back to its conventional multiplicative form. We will not be studying here the details of computing production function since there are a number of computer programs available for this purpose.

Types of Statistical Analyses

Once a functional form of a production function is chosen the next step is to select the type of statistical analysis to be used in its estimation. Generally, there are three types of statistical analyses used for estimation of a production function. These are: (a) time series analysis, (b) cross-section analysis and (c) engineering analysis.

- a) **Time series analysis:** The amount of various inputs used in various periods in the past and the amount of output produced in each period is called time series data. For example, we may obtain data concerning the amount of labour, the amount of capital, and the

amount of various raw materials used in the steel industry during each year from 1970 to 2000. On the basis of such data and information concerning the annual output of steel during 1970 to 2000, we may estimate the relationship between the amounts of the inputs and the resulting output, using regression techniques.

Analysis of time series data is appropriate for a single firm that has not undergone significant changes in technology during the time span analyzed. That is, we cannot use time series data for estimating the production function of a firm that has gone through significant technological changes. There are even more problems associated with the estimation a production function for an industry using time series data. For example, even if all firms have operated over the same time span, changes in capacity, inputs and outputs may have proceeded at a different pace for each firm. Thus, cross section data may be more appropriate.

- b) **Cross-section analysis:** The amount of inputs used and output produced in various firms or sectors of the industry at a given time is called cross- section data. For example, we may obtain data concerning the amount of labour, the amount of capital, and the amount of various raw materials used in various firms in the steel industry in the year 2020. On the basis of such data and information concerning the year 2020, output of each firm, we may use regression techniques to estimate the relationship between the amounts of the inputs and the resulting output.
- c) **Engineering analysis:** In this analysis we use technical information supplied by the engineer or the agricultural scientist. This analysis is undertaken when the above two types do not suffice. The data in this analysis is collected by experiment or from experience with day-to-day working of the technical process. There are advantages to be gained from approaching the measurement of the production function from this angle. Because the range of applicability of the data is known, and, unlike time- series and cross-section studies, we are not restricted to the narrow range of actual observations.

Limitations of Different Types of Statistical Analysis

Each of the methods discussed above has certain limitations.

1. Both time-series and cross-section analysis are restricted to a relatively narrow range of observed values. Extrapolation of the production function outside that range may be seriously misleading. For example, in a given case, marginal productivity might decrease

rapidly above 85% capacity utilization; the production function derived for values in the 70%-85% capacity utilization range would not show this.

2. Another limitation of time series analysis is the assumption that all observed values of the variables pertains to one and the same production function. In other words, a constant technology is assumed. In reality, most firms or industries, however, find better, faster, and/or cheaper ways of producing their output. As their technology changes, they are actually creating new production functions. One way of coping with such technological changes is to make it one of the independent variables.
3. Theoretically, the production function includes only efficient (least-cost) combinations of inputs. If measurements were to conform to this concept, any year in which the production was less than nominal would have to be excluded from the data. It is very difficult to find a time-series data, which satisfy technical efficiency criteria as a normal case.
4. Engineering data may overcome the limitations of time series data but mostly they concentrate on manufacturing activities. Engineering data do not tell us anything about the firm's marketing or financial activities, even though these activities may directly affect production.
5. In addition, there are both conceptual and statistical problems in measuring data on inputs and outputs.

It may be possible to measure output directly in physical units such as tons of coal, steel etc. In case more than one product is being produced, one may compute the weighted average of output, the weights being given by the cost of manufacturing these products. In a highly diversified manufacturing unit, there may be no alternative but to use the series of output values, corrected for changes in the price of products. One has also to choose between 'gross value' and 'net value'. It seems better to use "net value added" concept instead of output concept in estimating production function, particularly where raw-material intensity is high.

The data on labour is mostly available in the form of "number of workers employed" or "hours of labour employed". The 'number of workers' data should not be used because, it may not reflect underemployment of labour, and they may be occupied, but not productively employed. Even if we use 'man hours' data, it should be adjusted for efficiency factor. It is also not advisable that labour should be measured in monetary terms as given by expenditure on wages, bonus, etc.

The data on capital input has always posed serious problems. Net investment i.e. a change in the value of capital stock, is considered most appropriate.

Nevertheless, there are problems of measuring depreciation in fixed capital, changes in quality of fixed capital, changes in inventory valuation, changes in composition and productivity of working capital, etc.

Finally, when one attempts an econometric estimate of a production function, one has to overcome the standard problem of multi-collinearity among inputs, autocorrelation, homoscedasticity, etc.

7.8 MANAGERIAL USES OF PRODUCTION FUNCTION

There are several managerial uses of the production function. It can be used to compute the least-cost combination of inputs for a given output or to choose the input combination that yields the maximum level of output with a given level of cost. There are several feasible combinations of input factors and it is highly useful for decision-makers to find out the most appropriate among them. The production function is useful in deciding on the additional value of employing a variable input in the production process. So long as the marginal revenue productivity of a variable factor exceeds its price, it may be worthwhile to increase its use. The additional use of an input factor should be stopped when its marginal revenue productivity just equals its price. Production functions also aid long-run decision-making. If returns to scale are increasing, it will be worthwhile to increase production through a proportionate increase in all factors of production, provided, there is enough demand for the product. On the other hand, if returns to scale are decreasing, it may not be worthwhile to increase the production through a proportionate increase in all factors of production, even if there is enough demand for the product. However, it may be in the discretion of the producer to increase or decrease production in the presence of constant returns to scale, if there is enough demand for the product.

Activity 2

1. Can you list some more managerial uses of production function other than those given in section 7.8?

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

7.9 SUMMARY

A production function specifies the maximum output that can be produced with a given set of inputs. In order to achieve maximum profits the production manager has to use optimum input-output combination for a given cost. In this unit, we have shown how a production manager minimizes the cost for a given output in order to maximize the profit. Also, we have shown how to maximize the output at a given level of cost.

The law of diminishing marginal returns states that as equal increments of variable input are added to fixed input, a point will eventually be reached where corresponding increments to output begin to decline. We have also seen the relations between the marginal product, average product, and total product. There are three stages of production. Stage I is characterized by $MP > 0$ and $MP > AP$. Stage II is characterized by $MP > 0$ and $MP < AP$. Stage III is characterized by $MP < 0$. The economically meaningful range is Stage II. The production manager maximizes the profit at a point where the value of marginal product equals the price of the output.

A production isoquant consists of all the combinations of two inputs that will yield the same maximum output. The marginal rate of technical substitution is $\Delta K / \Delta L$, holding output constant. The law of diminishing marginal rate of substitution implies the rate at which one input can be substituted for another input, if output remains constant. An isocost line consists of all the combinations of inputs which have the same total cost. The absolute slope of the isocost line is the input price ratio. Returns to scale, a long run concept, involves the effect on output of changing all inputs by same proportion and in the same direction. Although, there are many different forms of production function we have discussed here only Cobb-Douglas production function.

7.10 SELF-ASSESSMENT QUESTIONS

1. Fill in the blanks in the following Table.

Units of labour	Total product	Average product	Marginal product
1		40	
2			48
3	138		
4		44	
5			24
6	210		
7		29	
8			-27

2. The marginal product of labour is known to be greater than the average product of labour at a given level of employment. Is the average product increasing or decreasing? Explain.
3. Explain the law of diminishing marginal returns and provide an example of the phenomenon.
4. Explain why a profit maximizing firm using only one variable input will produce in stage-II.
5. Explain why an A P curve and the corresponding MP curve must intersect at the maximum point on the A P curve
6. Explain why MP is greater than (less than) AP when AP is rising (falling).
7. Suppose a firm is currently using 500 labourers and 325 units of capital to produce its product. The wage rate is ₹25, and price of capital is ₹130. The last labourer adds 25 units of total output, while the last unit of capital adds 65 units to total output. Is the manager of this firm making the optimal input choice? Why or why not? If not, what should the manager do?

7.11 FURTHER READINGS

Adhikary, M. (1987). *Managerial Economics* (3rd ed.). Khosla Publishers, Delhi.

Maddala, G. S., & Miller, E. M. (1989). *Micro Economics: Theory and Applications*. McGraw-Hill, New York.

Maurice, S.C., Smithson, C.W., & Thomas, C. R. (2001). *Managerial Economics: Applied Microeconomics for Decision Making*. McGraw-Hill Publishing.

Mote, V. L., Paul, S., & Gupta, G. S. (2016). *Managerial Economics: Concepts and Cases*. Tata McGraw-Hill, New Delhi.

Dholakia, R., & Oza, A. N. (1996). *Microeconomics for Management Students*. Oxford University Press, Delhi.

UNIT 8 SHORT RUN COST ANALYSIS

Objectives

- After going through this unit, you should be able to:
- understand some of the cost concepts that are frequently used in the managerial decision-making process
- understand short run cost function
- understand applications of short run cost function in managerial decision making

Structure

- 8.1 Introduction
- 8.2 Actual Costs and Opportunity Costs
- 8.3 Explicit and Implicit Costs
- 8.4 Accounting Costs and Economic Costs
- 8.5 Direct Costs and Indirect Costs
- 8.6 Total Cost, Average Cost and Marginal Cost
- 8.7 Fixed and Variable Costs
- 8.8 Short-Run and Long-Run Costs
- 8.9 Short Run Cost Function
- 8.10 Applications of Short Run Cost Analysis
- 8.11 Summary
- 8.12 Self-Assessment Questions
- 8.13 Further Readings

8.1 INTRODUCTION

The analysis of cost is important in the study of managerial economics because it provides a basis for two important decisions made by managers: (a) whether to produce or not and (b) how much to produce when a decision is taken to produce. There are two types of cost analysis: **short run cost analysis & long run cost analysis**.

In this Unit, we shall discuss some important cost concepts that are relevant for managerial decisions. We analyze the basic differences between these cost concepts and also, examine how accountants and economists differ on treating different cost concepts. We shall discuss short run cost function and its applications in managerial decision making. The short run cost estimates a helpful to managers in arriving at the optimal mix of inputs to achieve a particular output target of a firm.

8.2 ACTUAL COSTS AND OPPORTUNITY COSTS

Actual costs are those costs, which a firm incurs while producing or acquiring a good or service like raw materials, labour, rent, etc. Suppose, we pay ₹ 150 per day to a worker whom we employ for 10 days, then the cost of labour is ₹ 1500. The economists called this cost as accounting costs because traditionally accountants have been primarily connected with collection of historical data (that is the costs actually incurred) in reporting a firm's financial position and in calculating its taxes. Sometimes the actual costs are also called **acquisition costs** or **outlay costs**.

On the other hand, **opportunity cost** is defined as the value of a resource in its next best use. For example, Mr. Ram is currently working with a firm and earning ₹ 5 lakhs per year. He decides to quit his job and start his own small business. Although, the accounting cost of Mr. Ram's labour to his own business is 0, the opportunity cost is ₹ 5 lakhs per year. Therefore, the opportunity cost is the earnings he foregoes by working for his own firm. One may ask you that whether this opportunity cost is really meaningful in the decision-making process. Opportunity cost can be similarly defined for other factors of production. For example, consider a firm that owns a building and therefore do not pay rent for office space. If the building was rented to others, the firm could have earned rent. The foregone rent is an opportunity cost of utilizing the office space and should be included as part of the cost of doing business. Sometimes these opportunity costs are called as alternative costs.

8.3 EXPLICIT AND IMPLICIT COSTS

Explicit costs are those costs that involve an actual payment to other parties. Therefore, an explicit cost is the monetary payment made by a firm for use of an input owned or controlled by others. Explicit costs are also referred to as accounting costs. For example, a firm pays ₹ 100 per day to a worker and engages 15 workers for 10 days, the explicit cost will be ₹ 15,000 incurred by the firm. Other types of explicit costs include purchase of raw materials, renting a building, amount spent on advertising etc.

On the other hand, **implicit costs** represent the value of foregone opportunities but do not involve an actual cash payment. Therefore, an implicit cost is the opportunity cost of using resources that are owned or controlled by the owners of the firm. Implicit costs are just as important as explicit costs but are sometimes neglected because they are not as obvious. The implicit cost is the foregone return, the owner of the firm could have received had they used their own resources in their best alternative use rather than using the resources for their own firm's production. For example, a manager who runs his own business foregoes the salary that could have been earned working for someone else as we have seen in our earlier example.

This implicit cost generally is not reflected in accounting statements, but rational decision-making requires that it be considered.

8.4 ACCOUNTING COSTS AND ECONOMIC COSTS

For a long time, there has been a considerable disagreement among economists and accountants on how costs should be treated. The reason for the difference of opinion is that the two groups want to use the cost data for dissimilar purposes. Accountants always have been concerned with firms' financial statements. Accountants tend to take a retrospective look at firm's finances because they keep trace of assets and liabilities and evaluate past performance. The **accounting costs** are useful for managing taxation needs as well as to calculate profit or loss of the firm. On the other hand, economists take forward-looking view of the firm. They are concerned with what cost is expected to be in the future and how the firm might be able to rearrange its resources to lower its costs and improve its profitability. They must therefore be concerned with opportunity cost. Since the only cost that matters for business decisions are the future costs, it is the economic costs that are used for decision-making. Accountants and economists both include explicit costs in their calculations. For accountants, explicit costs are important because they involve direct payments made by a firm. These explicit costs are also important for economists as well because the cost of wages and materials represent money that could be useful elsewhere.

Accountants and economists use the term 'profits' differently. Accounting profits are the firm's total revenue less its explicit costs. But economists define profits differently. Economic profits are total revenue less all costs (explicit and implicit costs). The economist takes into account the implicit costs (including a normal profit) in addition to explicit costs in order to retain resources in a given line of production. Therefore, when an economist says that a firm is just covering its costs, it is meant that all explicit and implicit costs are being met, and that, the entrepreneur is receiving a return just large enough to retain his/her talents in the present line of production. If a firm's total receipts exceed all its economic costs, the residual accruing to the entrepreneur is called an economic profit, or pure profit.

Example of Economic Profit and Accounting Profit

Mr. Raj is a small storeowner. He has invested ₹ 2 lakhs as equity in the store and inventory. His annual turnover is ₹ 8 lakhs, from which he must deduct the cost of goods sold, salaries of hired staff, and depreciation of equipment and building to arrive at annual profit of the store. He asked help of a friend who is an accountant by profession to prepare annual income statement. The accountant reported the profit to be ₹ 1.5 lakhs. Mr. Raj could not believe this and asked the help of another friend who is an economist by profession. The economist told him that the actual profit was

only ₹ 75,000 and not ₹ 1.5 lakhs. The economist found that the accountant had underestimated the costs by not including the implicit costs of time spent as Manager by Mr. Raj in the business and interest on owner's equity. The two income statements are shown below:

Income statement prepared by accountant			Income statement prepared by economist		
	₹	₹		₹	₹
Sales		8,00,000	Sales		8,00,000
<i>Explicit costs</i>			<i>Explicit costs</i>		
Cost of goods sold	6,00,000		Cost of goods sold	6,00,000	
Salaries	40,000		Salaries	40,000	
Depreciation	10,000	6,50,000	Depreciation	10,000	6,50,000
			<i>Implicit costs</i>		
			Salary to owner Manager	50,000	
			Interest on owners' equity	25,000	75,000
Accounting profit		1,50,000	Economic profit		75,000

Controllable and Non-Controllable costs

Controllable costs are those which are capable of being controlled or regulated by executive vigilance and, therefore, can be used for assessing executive efficiency. Non-controllable costs are those, which cannot be subjected to administrative control and supervision. Most of the costs are controllable, except, of course, those due to obsolescence and depreciation. The level at which such control can be exercised, however, differs: some costs (like, capital costs) are not controllable at factory's shop level, but inventory costs can be controlled at the shop level.

Out-of-pocket costs and Book costs

Out of pocket costs are those costs that involve current cash payments to outsiders. For example, wages and salaries paid to the employees are out-of-pocket costs. Other examples of out-of-pocket costs are payment of rent, interest, transport charges, etc. On the other hand, **book costs** are those business costs, which do not involve any cash payments but for them a provision is made in the books of account to include them in profit and loss accounts and take tax advantages. For example, salary of owner manager, if not paid, is a book cost. The interest cost of owner's own fund and depreciation cost are other examples of book cost. The out-of-pocket costs are also called explicit costs and correspondingly book costs are called implicit or imputed costs.

Past and Future costs

Past costs are actual costs incurred in the past and they are always contained in the income statements. Their measurement is essentially a record keeping activity. These costs can only be observed and evaluated in retrospect. If they are regarded as excessive, management can indulge in post-mortem checks to find out the factors responsible for the excessive costs, if any and take corrective actions for future. Past costs serve as the basis for projecting future costs.

Future costs are those costs that are likely to be incurred in future periods. Since the future is uncertain, these costs have to be estimated and cannot be expected to be absolutely correct figures. In periods of inflation and deflation, the two cost concepts differ significantly. Managerial decisions are forward looking and therefore they require estimates of future costs. Unlike past costs, future costs are subject to management control and they can be planned or avoided. Management needs to estimate future costs for a variety of reasons such as expense control pricing, projecting future profits and capital budgeting decisions. When historical costs are used instead of explicit projections, the assumption is made that future costs will be the same as past costs. In periods of significant price variations, such an assumption may lead to wrong managerial decisions.

Historical and Replacement costs

The historical cost of an asset is the actual cost incurred at the time; the asset was originally acquired. In contrast to this, replacement cost is the cost, which will have to be incurred if that asset is purchased now. The difference between the historical and replacement costs results from price changes over time. Suppose a machine was acquired for ₹50,000 in the year 2015 and the same machine can be acquired for ₹ 1,20,000 in the year 2021. Here ₹ 50,000 is the historical or original cost of the machine and ₹ 1,20,000 is its replacement cost. The difference of ₹70,000 between the two costs has resulted because of the price change of the machine during the period. In the conventional financial accounts, the value of assets is shown at their historical costs. But for decision-making, firms should try to adjust historical costs to reflect price level changes. If the price of the asset does not change over time, the historical cost will be the same as the replacement cost. If the price raises the replacement cost will exceed historical cost and vice versa. During periods of substantial price variations, historical costs are poor indicators of actual costs.

Private Costs and Social Costs

A further distinction that is useful to make - especially in the public sector - is between private and social costs. **Private costs** are those that accrue directly to the individuals or firms engaged in relevant activity. Social costs, on the

other hand, are passed on to persons not involved in the activity in any direct way (i.e., they are passed on to society at large). Consider the case of a manufacturer located on the bank of a river who dumps the waste into water rather than disposing it of in some other manner. While the private cost to the firm of dumping is zero, it is definitely harmful to the society. It affects adversely the people located down current and incur higher costs in terms of treating the water for their use, or having to travel a great deal to fetch potable water. If these external costs were included in the production costs of a producing firm, a true picture of real or social costs of the output would be obtained. Ignoring external costs may lead to an inefficient and undesirable allocation of resources in society.

Relevant Costs and Irrelevant Costs

The **relevant costs** for decision-making purposes are those costs, which are incurred as a result of the decision under consideration. The relevant costs are also referred to as the incremental costs. Costs that have been incurred already and costs that will be incurred in the future, regardless of the present decision are **irrelevant costs** as far as the current decision problem is concerned.

There are three main categories of relevant or incremental costs. These are the present-period explicit costs, the opportunity costs implicitly involved in the decision, and the future cost implications that flow from the decision. For example, direct labour and material costs, and changes in the variable overhead costs are the natural consequences of a decision to increase the output level.

Also, if there is any expenditure on capital equipments incurred as a result of such a decision, it should be included in full, notwithstanding that the equipment may have a useful life remaining after the present decision has been carried out. Thus, the incremental costs of a decision to increase output level will include all present-period explicit costs, which will be incurred as a consequence of this decision. It will exclude any present-period explicit cost that will be incurred regardless of the present decision.

The opportunity cost of a resource under use, as discussed earlier, becomes a relevant cost while arriving at the economic profit of the firm. Many decisions will have implications for future costs, both explicit and implicit. If a firm expects to incur some costs in future as a consequence of the present analysis, such future costs should be included in the present value terms if known for certain.

Sunk Costs and Incremental Costs

Sunk costs are expenditures that have been made in the past or must be paid in the future as part of contractual agreement or previous decision. For example, the money already paid for machinery, equipment, inventory and future rental payments on a warehouse that must be paid as part of long-term lease agreement are sunk costs. In general, sunk costs are not relevant to

economic decisions. For example, the purchase of specialized equipment designed to order for a plant. We assume that the equipment can be used to do only what it was originally designed for and cannot be converted for alternative use. The expenditure on this equipment is a sunk cost. Also, because this equipment has no alternative use its opportunity cost is zero and, hence, sunk costs are not relevant to economic decisions. Sometimes the sunk costs are also called as non-avoidable or non-escapable costs.

On the other hand, **incremental cost** refers to total additional cost of implementing a managerial decision. Change in product line, change in output level, adding or replacing a machine, changing distribution channels etc. are examples of incremental costs. Sometimes incremental costs are also called as avoidable or escapable costs. Moreover, since incremental costs may also be regarded as the difference in total costs resulting from a contemplated change, they are also called **differential costs**.

As stated earlier sunk costs are irrelevant for decision making, as they do not vary with the changes contemplated for future by the management. It is the incremental costs, which are important for decision-making purpose.

Activity 1

1. A graduate from a business school decides to open business and devote his full time to its management. What cost would you assign to his time? Is this implicit or explicit?
.....
.....
.....
.....
.....
2. Why is the historic cost of inventory or capital equipment irrelevant for managerial decision making?
.....
.....
.....
.....
.....
3. Comment on the nature of costs involved in depreciation from both economic and accounting approaches.
.....
.....
.....
.....
.....

8.5 DIRECT COSTS AND INDIRECT COSTS

There are some costs, which can be directly attributed to production of a given product. The use of raw material, labour input, and machine time involved in the production of each unit can usually be determined. On the other hand, there are certain costs like stationery and other office and administrative expenses, electricity charges, depreciation of plant and buildings, and other such expenses that cannot easily and accurately be separated and attributed to individual units of production, except on arbitrary basis. When referring to the separable costs of first category accountants call them the direct, or prime costs per unit. The accountants refer to the joint costs of the second category as indirect or overhead costs. **Direct and indirect costs** are not exactly synonymous to what economists refer to as variable costs and fixed costs. The criterion used by the economist to divide cost into either fixed or variable is whether or not the cost varies with the level of output, whereas the accountant divides the cost on the basis of whether or not the cost is separable with respect to the production of individual output units. The accounting statements often divide overhead expenses into 'variable overhead' and 'fixed overhead' categories. If the variable overhead expenses per unit are added to the direct cost per unit, we arrive at what economists call as average variable cost.

Separable Costs and Common Costs

Costs can also be classified on the basis of their traceability. The costs that can be easily attributed to a product, a division, or a process are called **separable costs**. On the other hand, **common costs** are those, which cannot be traced to any one unit of operation. For example, in a multiple product firm the cost of raw material may be separable (traceable) product-wise but electricity charges may not be separable product-wise. In a University the salary of a Vice-Chancellor is not separable department-wise but the salary of teachers can be separable department-wise. The separable and common costs are also referred to as direct and indirect costs respectively. The distinction between direct and indirect costs is of particular significance in a multi-product firm for setting up economic prices for different products.

8.6 TOTAL COST, AVERAGE COST & MARGINAL COST

Total cost (TC) of a firm is the sum-total of all the explicit and implicit expenditures incurred for producing a given level of output. It represents the money value of the total resources required for production of goods and services. For example, a shoe-maker's total cost will include the amount s/he spends on leather, thread, rent for her/his workshop, interest on borrowed capital, wages and salaries of employees, etc., and the amount s/he charges

for her/his services and funds invested in the business. The total cost concept is useful in break-even analysis and finding out whether a firm is making profit or not.

Average cost (AC) is the cost per unit of output. That is, average cost equals the total cost divided by the number of units produced (N). If $TC = ₹500$ and $N = 50$ then $AC = ₹10$. The average cost concept is significant for calculating the per unit profit.

Marginal cost (MC) refers to the change in total cost associated with a one-unit change in output. Marginal cost (MC) is the extra cost of producing one additional unit. At a given level of output, one examines the additional costs being incurred in producing one extra unit and this yields the marginal cost. For example, if TC of producing 100 units is ₹10,000 and the TC of producing 101 units is ₹10,050, then MC at $N = 101$ equals ₹50. This cost concept is significant to short-term decisions about profit maximizing rates of output. For example, in an automobile manufacturing plant, the marginal cost of making one additional car per production period would be the labour, material, and energy costs directly associated with that extra car. Marginal cost is that sub category of incremental cost in the sense that incremental cost may include both fixed costs and marginal costs.

However, when production is not conceived in small units, management will be interested in incremental cost instead of marginal cost. For example, if a firm produces 5000 units of TV sets, it may not be possible to determine the change in cost involved in producing 5001 units of TV sets. This difficulty can be resolved by taking units to significant size. For example, if the TV sets produced is measured to hundreds of units and total cost (TC) of producing the current level of three hundred TV sets is ₹15,00,000 and the firm decides to increase the production to four hundred TV sets and estimates the TC as ₹18,00,000, then the incremental cost of producing one hundred TV sets (above the present production level of three hundred units) is ₹3,00,000.

The **marginal and incremental cost concepts** are needed in deciding whether a firm needs to expand its production or not. In fact, the relevant costs to be considered will depend upon the situation or production problem faced by the manager.

8.7 FIXED COSTS AND VARIABLE COSTS

Fixed costs are that part of the total cost of the firm which does not change with output. Expenditures on depreciation, rent of land and buildings, property taxes, and interest payment on bonds are examples of fixed costs. Given a capacity, fixed costs remain the same irrespective of actual output. **Variable costs**, on the other hand, change with changes in output. Examples of variable costs are wages and expenses on raw material.

However, it is not very easy to classify all costs into fixed and variable. There are some costs, which fall between these extremes. They are called **semi-variable costs**. They are neither perfectly variable nor absolutely fixed in relation to changes in output. For example, part of the depreciation charges is fixed and part variable. However, it is very difficult to determine how much of depreciation cost is due to the technical obsolescence of assets and hence fixed cost, and how much is due to the use of equipments and hence variable cost. Nevertheless, it does not mean that it is not useful to classify costs in to fixed and variable. This distinction is of great value in break-even analysis and pricing decisions. For decision-making purposes, in general, it is the variable cost, which is relevant and not the fixed cost.

To an economist the fixed costs are overhead costs and to an accountant these are indirect costs. When the output goes up, the fixed cost per unit of output comes down, as the total fixed cost is divided between larger units of output.

8.8 SHORT RUN AND LONG RUN COSTS

The **short run** is defined as a period in which the supply of at least one element of the inputs cannot be changed. To illustrate, certain inputs like machinery, buildings, etc., cannot be changed by the firm whenever it so desires. It takes time to replace, add or dismantle them. Short-run costs are the costs that can vary with the degree of utilization of plant and other fixed factors. In other words, these costs relate to the variation in output, given plant capacity. Short-run costs are, therefore, of two types: fixed costs and variable costs. In the short-run, fixed costs remain unchanged while variable costs fluctuate with output.

Long run, on the other hand, is defined as a period in which all inputs are changed with changes in output. In other words, it is that time-span in which all adjustments and changes are possible to realize. Long-run costs are costs that can vary with the size of plant and with other facilities normally regarded as fixed in the short-run. In fact, in the long-run there are no fixed inputs and therefore no fixed costs, i.e., all costs are variable. Thus, in the short run, some inputs are fixed (like installed capacity) while others are variable (like the level of capacity utilization); but in the long run all inputs, including the size of the plant, are variable.

Both short-run and long-run costs are useful in decision-making. Short-run cost is relevant when a firm has to decide whether or not to produce and if a decision is taken to produce then how much more or less to produce with a given plant size. If the firm is considering an increase in plant size, it must examine the long-run cost of expansion. Long-run cost analysis is useful in investment decisions.

Activity 2

1. Classify different cost concepts into:

a. useful for decision making

.....

.....

.....

.....

.....

b. Not useful for decision making

.....

.....

.....

.....

.....

2. Give specific example of:
- a. Direct costs:.....
 - b. Indirect costs:
 - c. Sunk costs:
 - d. Traceable costs:
 - e. Common costs :
 - f. Implicit costs:
 - g. Social costs:
 - h. Non-controllable costs :

3. Give examples to distinguish between ‘fixed overheads’ and ‘variable overheads’.

.....

.....

.....

.....

.....

4. Can all ‘direct costs’ be treated as ‘variable costs’ Explain?

.....

.....

.....

.....

.....

8.9 SHORT RUN COST FUNCTIONS

In above sections, we have distinguished between the short run and the long run. We also distinguished between fixed costs and variable costs. The distinction between fixed and variable costs is of great significance to the business manager. Variable costs are those costs, which the business manager can control or alter in the short run by changing levels of production. On the other hand, fixed costs are clearly beyond business manager's control, such costs are incurred in the short run and must be paid regardless of output.

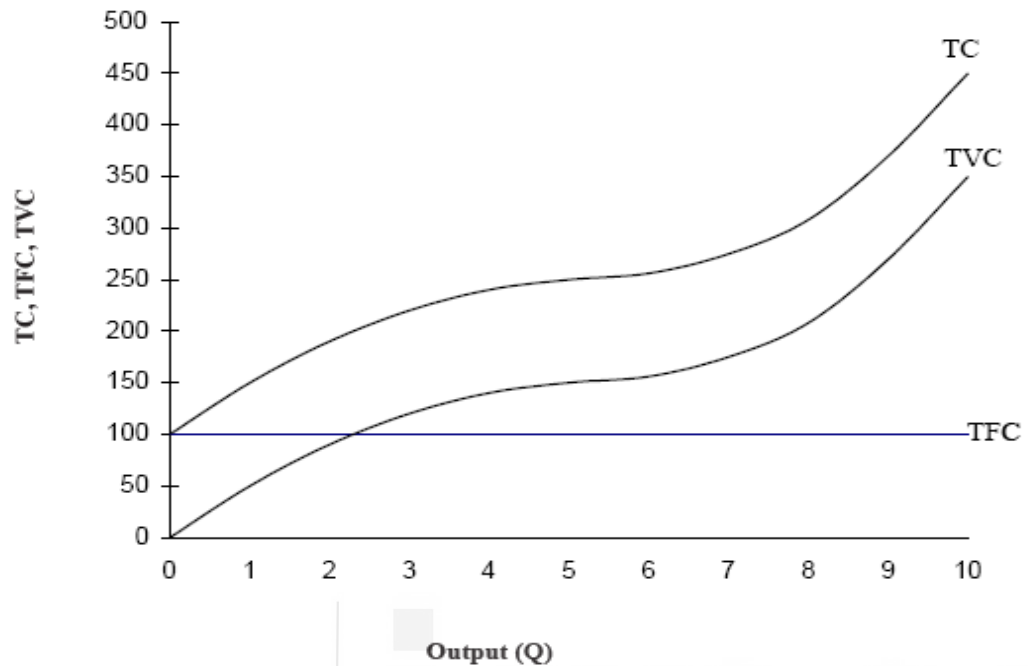
Total Costs

Three concepts of total cost in the short run must be considered: total fixed cost (TFC), total variable cost (TVC), and total cost (TC). **Total fixed costs** are the total costs per period of time incurred by the firm for fixed inputs. Since the amount of the fixed inputs is fixed, the total fixed cost will be the same regardless of the firm's output rate. Table 8.1 shows the costs of a firm in the short run. According to this table, the firm's total fixed costs are ₹100. The firm's total fixed cost function is shown graphically in Figure 8.1.

Table 8.1: A Firm's Short Run Costs (in ₹)

Q	TFC	TVC	TC	MC	AFC	AVC	ATC
0	100	0	100				
1	100	50	150	50	100.0	50	150
2	100	90	190	40	50.0	45	95.0
3	100	120	220	30	33.3	40	73.3
4	100	140	240	20	25.0	35	60.0
5	100	150	250	10	20.0	30	50.0
6	100	156	256	6	16.7	26	42.7
7	100	175	275	19	14.3	25	39.3
8	100	208	308	33	12.5	26	38.5
9	100	270	370	62	11.1	30	41.1
10	100	350	450	80	10.0	35	45.0

Figure 8.1: Total Cost Curves



Total variable costs are the total costs incurred by the firm for variable inputs. To obtain total variable cost we must know the price of the variable inputs. Suppose if we have two variable inputs viz. labour (V_1) and raw material (V_2) and the corresponding prices of these inputs are P_1 and P_2 , then the total variable cost (TVC) = $P_1 \times V_1 + P_2 \times V_2$. They go up as the firm's output rises, since higher output rates require higher variable input rates, which mean bigger variable costs. The firm's total variable cost function corresponding to the data given in Table 8.1 is shown graphically in Figure 8.1.

Finally, *total costs* are the sum of total fixed costs and total variable costs. To derive the total cost column in Table 8.1, add total fixed cost and total variable cost at each output. The firm's total cost function corresponding to the data given in Table 8.1 is shown graphically in Figure 8.1. Since total fixed costs are constant, the total fixed cost curve is simply a horizontal line at ₹100. And because total cost is the sum of total variable costs and total fixed costs, the total cost curve has the same shape as the total variable cost curve but lies above it by a vertical distance of ₹100.

Corresponding to our discussion above we can define the following for the short run:

$$TC = TFC + TVC$$

where,

TC = total cost

TFC = total fixed costs

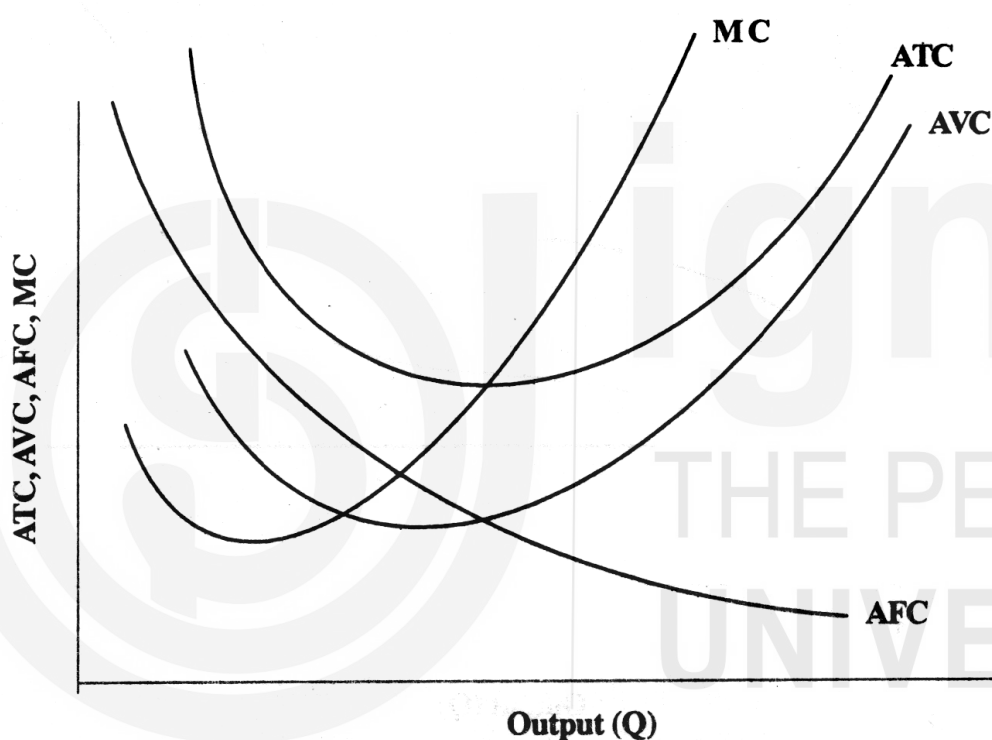
TVC = total variable costs

Average Fixed Costs

While the total cost functions are of great importance, managers must be interested as well in the average cost functions and the marginal cost function as well. There are three average cost concepts corresponding to the three total cost concepts. These are average fixed cost (AFC), average variable cost (AVC), and average total cost (ATC). Figure 8.2 shows typical average fixed cost function graphically. *Average fixed cost* is the total fixed cost divided by output. Average fixed cost declines as output (Q) increases. Thus, we can write average fixed cost as:

$$AFC = TFC/Q$$

Figure 8.2: Short Run Average and Marginal Cost Curves



Average Variable Costs

Average variable cost is the total variable cost divided by output. Figure 8.2 shows the average variable cost function graphically. At first, output increases resulting in decrease in average variable cost, but beyond a point, they result in higher average variable cost.

$$AVC = \frac{TVC}{Q}$$

where,

Q = Output

TVC = total variable costs

AVC = average variable costs

Average Total Cost

Average total cost (ATC) is the sum of the average fixed cost and average variable cost. In other words, ATC is total cost divided by output. Thus,

$$ATC = AFC + AVC = \frac{TC}{Q}$$

Figure 8.2 shows the average total cost function graphically. Since ATC is sum of the AFC and AVC, ATC curve always exceeds AVC curve. Also, since AFC falls as output increases, AVC and ATC get closer as output rises. Note that ATC curve is nearer the AFC curve at initial levels of output, but is nearer the AVC curve at later levels of output. This indicates that at lower levels of output fixed costs are more important part of the total cost, while at higher levels of output the variable element of cost becomes more important.

Marginal Cost

Marginal cost (MC) is the addition to either total cost or total variable cost resulting from the addition of one unit of output. Thus,

$$MC = \frac{\Delta TC}{\Delta Q} = \frac{\Delta TVC}{\Delta Q}$$

where,

MC = marginal cost

ΔQ = change in output

ΔTC = change in total cost due to change in output

ΔTVC = change in total variable cost due to change in output

The two definitions are the same because, when output increases, total cost increases by the same amount as the increase in total variable cost (since fixed cost remains constant). Figure 8.2 shows the marginal cost function graphically. At low output levels, marginal cost may decrease with increase in output, but after reaching a minimum, it goes up with further increase in output. The reason for this behavior is found in diminishing marginal returns.

The marginal cost concept is very crucial from the manager's point of view. Marginal cost is a strategic concept because it designates those costs over which the firm has the most direct control. More specifically, MC indicates those costs which are incurred in the production of the last unit of output and therefore, also the cost which can be "saved" by reducing total output by the last unit. Average cost figures do not provide this information. A firm's decisions as to what output level to produce is largely influenced by its marginal cost. When coupled with marginal revenue, which indicates the change in revenue from one more or one less unit of output, marginal cost allows a firm to determine whether it is profitable to expand or contract its level of production.

The relationships between the various average and marginal cost curves are illustrated in Figure 8.2. The figure shows typical AFC, AVC, ATC, and MC curves but is not drawn to scale for the data given in Table 8.1. The MC cuts both AVC and ATC at their minimum. When both the MC and AVC are falling, AVC will fall at a slower rate. When both the MC and AVC are rising, MC will rise at a faster rate. As a result, MC will attain its minimum before the AVC. In other words, when MC is less than AVC, the AVC will fall, and when MC exceeds AVC, AVC will rise. This means that as long as MC lies below AVC, the latter will fall and where MC is above AVC, AVC will rise.

Therefore, at the point of intersection where $MC = AVC$, AVC has just ceased to fall and attained its minimum, but has not yet begun to rise. Similarly, the MC curve cuts the ATC curve at the latter's minimum point. This is because MC can be defined as the addition either to TC or TVC resulting from one more unit of output. However, no such relationship exists between MC and AFC, because the two are not related; MC by definition includes only those costs which change with output, and FC by definition is independent of output.

Relationship between Average Product and Marginal Product, and Average Variable Cost and Marginal Cost

There is a straightforward relationship between factor productivity and output costs. To see this, let us consider a single variable factor L say labour. All other inputs are fixed. AP and MP will denote the average and marginal products of labour, respectively. If W is the wage rate and L is the quantity of labour, then

$$TVC = W * L$$

Hence, if Q is the output,

$$AVC = \frac{TVC}{Q} = W \left\{ \frac{L}{Q} \right\}$$

Consequently, since Q/L is the average product (AP), $AVC = W/AP$

Also, $\Delta TVC = W * \Delta L$ (W does not change and is assumed to be given.).

Dividing by ΔQ we get

$$MC = \frac{\Delta TVC}{\Delta Q} = W \left\{ \frac{\Delta L}{\Delta Q} \right\}$$

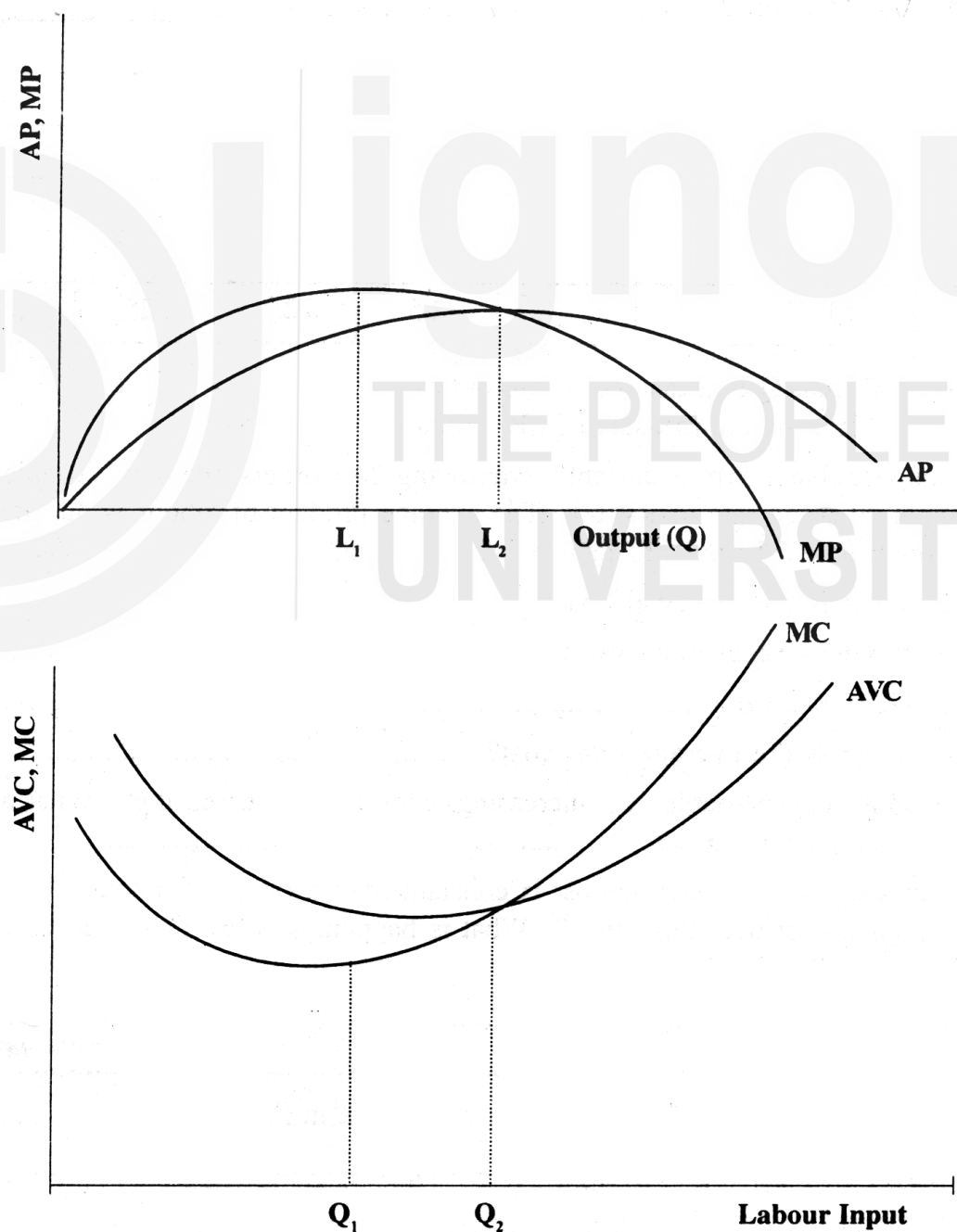
But, marginal product (MP) = $\Delta Q / \Delta L$. Hence, $MC = W/MP$

Figure 8.3 shows the relationship between average product and marginal product, and average variable cost and marginal cost. The relationship $AVC = W/AP$ shows that AVC is at a minimum when AP is at maximum.

Similarly, the relationship $MC = W/MP$ shows that MC is at a minimum when MP is at a maximum. Also, when AP is at a maximum, $AP = MP$. Hence, when AVC is at a minimum, $AVC = MC$. It is clearly shown that when MP is rising, MC is falling. And when MP is falling, MC is rising.

The relevant costs to be considered for decision-making will differ from one situation to the other depending on the problem faced by the manager. In general, the TC concept is quite useful in finding out the breakeven quantity of output. The TC concept is also used to find out whether firm is making profits or not. The AC concept is important for calculating the per unit profit of a business firm. The MC concept is essential to decide whether a firm should expand its production or not.

Figure 8.3: Relationship between AP and MP, AVC and MC



Activity 3

Short Run Cost Analysis

1. Fill in the blanks in the Table below:

Q	TFC	TVC	TC	AFC	AVC	ATC	MC
1.	50		55				
2.	50	8		25			
3.	50		60.5				
4.		13					
5.	50		65				
6.	50	18			3	11.3	3
7.	50		72.5				
8.	50	28					
9.			86				
10.	50	45		5		9.5	9
11.	50	54.5		4.5		9.5	9.5
12.	50	65.2					
13.	50		130				
14.	50	99.1					
15.	50		174.75				
16.	50	162					
17.	50		259.25				
18.		269.5					
19.	50		399				
20.	50	450		2.5	22.5	25	101

Note: Output Q is measured in '000 units

All costs are measured in ₹ '000

2. Suppose that a firm is currently employing 20 workers, the only variable input, at wage rate of ₹ 60. The average product of labour is 30, the last worker added 12 units to total output, and total fixed cost is ₹ 3600.

a. What is the marginal cost?

.....

b. What is the average variable cost?

.....

c. How much output is being produced?

.....

d. What is the average total cost?

.....

e. Is average variable cost increasing, constant, or decreasing? What about average total cost?

.....

3. Suppose average variable cost is constant over a range of output. What is marginal cost over this range? What is happening to average total cost over this range?

.....

.....

.....

.....

.....

8.10 APPLICATION OF SHORT RUN COST ANALYSIS

In the previous sections of this unit, we discussed total, marginal, and average cost curves for short run. The relationships between these cost curves have a very wide range of applications for managerial use. Here we will discuss a few applications of these concepts.

Determining Optimum Output Level

Earlier we have seen that the optimum output level is the point where average cost is minimum. In other words, the optimum output level is the point where average cost equals marginal cost. Consider the following example.

$$TC = 128 + 6Q + 2Q^2$$

This is a short run total cost function since there is a fixed cost (TFC = 128).

$$AC = (TC/Q) = \frac{128}{Q} + 6 + 2Q$$

$$\frac{d(AC)}{dQ} = -\frac{128}{Q^2} + 2 = 0$$

$$2Q^2 = 128$$

$$Q^2 = 64$$

$$Q = 8$$

or

$$MC = \frac{d(TC)}{dQ} = 6 + 4Q = 0$$

Setting $AC = MC$

$$\frac{128}{Q} + 6 + 2Q = 6 + 4Q$$

$$\frac{128}{Q} - 2Q = 0$$

$$2Q^2 = 128$$

$$Q^2 = 64$$

$$Q = 8$$

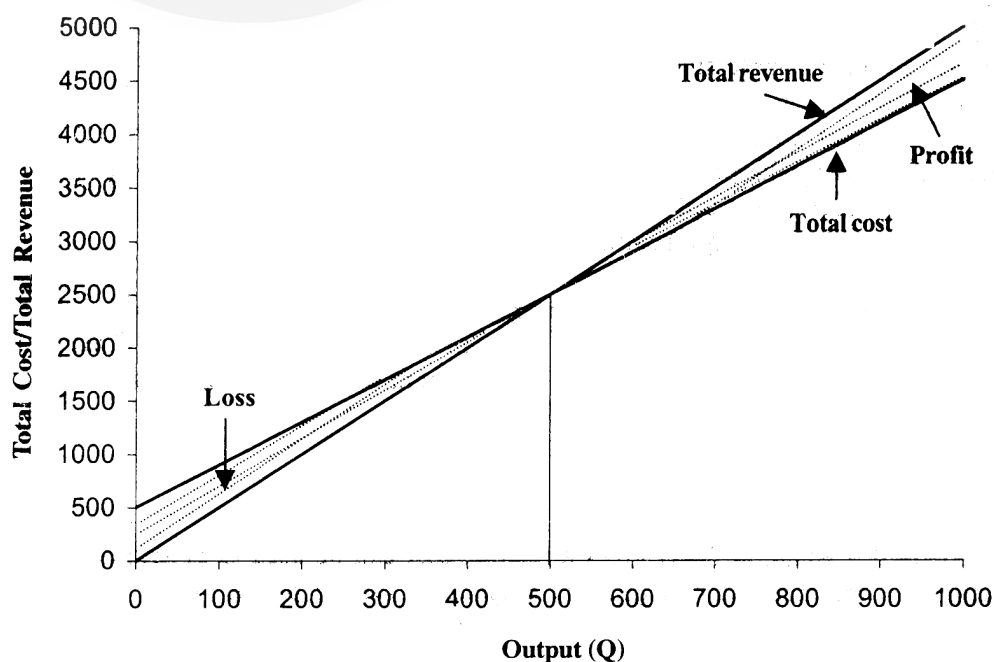
Thus $Q = 8$ and is the optimum level of output in the short run.

Breakeven Output Level

An analytical tool frequently employed by managerial economists is the breakeven chart, an important application of cost functions. The breakeven chart illustrates at what level of output in the short run, the total revenue just covers total costs. Generally, a breakeven chart assumes that the firm's average variable costs are constant in the relevant output range; hence, the firm's total cost function is assumed to be a straight line. Since variable cost is constant, the marginal cost is also constant and equals to average variable cost.

Figure 8.4 shows the breakeven chart of a firm. Here, it is assumed that the price of the product will not be affected by the quantity of sales. Therefore, the total revenue is proportional to output. Consequently, the total revenue curve is a straight line through the origin. The firm's fixed cost is ₹500, variable cost per unit is ₹4 and the unit sales price of output is ₹5. The breakeven chart, which combines the total cost function and the total revenue curve, shows profit or loss resulting from each sales level. For example, Figure 8.4 shows that if the firm sells 200 units of output it will make a loss of ₹300. The chart also shows the breakeven point, the output level that must be reached if the firm is to avoid losses. It can be seen from the figure; the breakeven point is 500 units of output. Beyond 500 units of output the firm makes profit.

Figure 8.4 Breakeven Chart



Breakeven charts are used extensively for managerial decision process. Under right conditions, breakeven charts can produce useful projections of the effect of the output rate on costs, revenue and profits. For example, a firm may use breakeven chart to determine the effect of projected decline in sales or profits. On the other hand, the firm may use it to determine how many units of a particular product it must sell in order to breakeven or to make a particular level of profit. However, breakeven charts must be used with caution, since the assumptions underlying them, sometimes, may not be appropriate. If the product price is highly variable or if costs are difficult to predict, the estimated total cost function and revenue curves may be subject to these errors.

We can analyze the breakeven output with familiar algebraic equations.

$$TR = P * Q$$

$$TC = TFC + AVC \times Q \left\{ AVC = \frac{TVC}{Q} \text{ so, } TVC = AVC * Q \right\}$$

At breakeven point, $TR = TC$

$$P * Q = TFC + (AVC * Q)$$

$$Q = \frac{TFC}{P - AVC} = \frac{\text{Total Fixed cost}}{\text{Price} - \text{Variable cost per unit}}$$

Here Q stands for breakeven volume of output. Multiplying Q with price (P) we get the breakeven value of output. In the case of our example given in Figure 8.4, $TFC = ₹500$, $P = ₹5$ and $AVC = ₹4$. Consequently,

$$Q = \frac{500}{5 - 4} = \frac{500}{1} = 500$$

Therefore, the breakeven output (Q) will be 500 units. Similarly, the breakeven output value will be ₹2500 ($P \times Q = ₹5 \times 500$).

Profit Contribution Analysis

In making short run decisions, firms often find it useful to carry out *profit contribution analysis*. The profit contribution is the difference between price and average variable cost ($P - AVC$). That is, revenue on the sale of a unit of output after variable costs are covered represents a contribution towards profit. In our example since price is ₹5 and average variable cost is ₹4, the profit contribution per unit of output will be ₹1 ($₹5 - ₹4$). At low rates of output the firm may be losing money because fixed costs have not yet been covered by the profit contribution. Thus, at these low rates of output, profit contribution is used to cover fixed costs. After fixed costs are covered, the firm will be earning a profit.

A manager wants to know the output rate necessary to cover all fixed costs and to earn a 'required' profit (π_R). Assume that both price and AVC are

constant. Profit is equal to revenue less the sum of total variable costs and fixed costs. Thus

$$\pi_R = TR - TC$$

$$\pi_R = (P * Q) - \{TFC + (AVC * Q)\}$$

Solving this equation for Q gives a relation that can be used to determine the rate of output necessary to generate a specified rate of profit. Thus

$$\pi_R + TFC = (P * Q) - (AVC * Q)$$

$$\pi_R + TFC = Q (P - AVC)$$

$$Q = \frac{\pi_R + TFC}{P - AVC}$$

To illustrate how profit contribution analysis can be used, suppose that the firm in our example (where $TFC = ₹ 500$, $P = ₹5$ and $AVC = ₹4$) wants to determine how many units of output it will have to produce and sell to earn a profit of ₹10, 000. To generate this profit, an output rate of 10,500 units is required; that is,

$$Q = \frac{10,000 + 500}{5 - 4} = 10500$$

Operating Leverage

Managers must make comparisons among alternative systems of production. Should one type of plant be replaced by another? Breakeven analysis can be extended to help make such comparisons more effective. Consider the *degree of operating leverage* (E_π), which is defined as the percentage change in profit resulting from a 1% change in the number of units of product sold. Thus,

$$E_\pi = \frac{\% \text{ change in profit}}{\% \text{ change in output sold}}$$

$$= \frac{(\Delta\pi / \pi)\Delta\pi}{\Delta Q / Q} = \frac{Q}{\Delta Q} * \frac{d\pi}{\pi} \text{ or } \frac{Q}{dQ} * \frac{d\pi}{\pi}$$

If the price of output is constant regardless of the rate of output, the change in degree of operating leverage depends on three variables: the rate of output, the level of fixed costs, and variable cost per unit of output. This can be seen by substituting the above equation for profit with

$$\pi = TR - (TVC + TFC)$$

$$\pi = P * Q - (AVC * Q) - TFC$$

$$\text{and change in profit } \Delta\pi = P * \Delta Q - (AVC) * \Delta Q$$

Therefore, the degree of operating leverage will be

$$E_\pi = \frac{[P * \Delta Q - (AVC) * \Delta Q] / [P * Q - (AVC) * Q - TFC]}{\Delta Q / Q}$$

On Simplification,

$$E_{\pi} = \frac{(P - AVQ)}{Q(P - AVQ) - TFC}$$

Example: Consider three firms I, II and III having the following fixed costs, average variable costs and price of the product.

Firm	Fixed Cost (₹)	Average variable Cost (₹)	Price of the product (₹)
Firm-I	1,00,000	2	5
Firm-II	60,000	3	5
Firm-III	26,650	4	5

Firm-I has more fixed cost than firm-II, and firm-III. However, Firm-I has less average costs than firm-II, and firm-III. Essentially, firm-I has substituted capital (fixed costs) for labour and materials (variable costs) with the introduction of more mechanized machines. On the other hand, firm-III has less fixed costs and more average variable costs when compared to other two plants because firm-III has less mechanized machines. The firm-II occupies middle position in terms of fixed costs and average variable costs.

In comparing these plants, we use the degree of operating leverage. Suppose for all the three plants $Q=40,000$.

$$\text{For firm - I, } E_{\pi} = \frac{40000 (5-2)}{40000 (5-2) - 100000} = 6$$

$$\text{For firm - II, } E_{\pi} = \frac{40000 (5-3)}{40000 (5-2) - 60000} = 4$$

$$\text{For firm - III, } E_{\pi} = \frac{40000 (5-4)}{40000 (5-4) - 26,650} = 3$$

Thus, a 1% increase in sales volume results in a 6% increase in profit at firm-I, a 4% profit at firm-II, and 3% profit at firm-III. This means firm-I's profits are more sensitive to changes in sales volume than firm-II and firm-III and firm-II's profits are more sensitive to changes in sales volume than firm-III.

Activity 4

- Speed-Marine Co. builds motorboat engines. They recently estimated their total costs and total revenue as:

$$TC = 80,000 - 600Q + 2Q^2 \quad TR = 400Q - Q^2$$

Where TC is total cost, TR is total revenue, and Q is the number of engines produced each year.

Where TC is total cost, TR is total revenue, and Q is the number of engines produced each year.

- a. At what level of production will the company breakeven? How many engines should be produced to maximize profit?

.....

.....

.....

.....

2. Give $TC=6Q + 2Q^2 - Q^3$, find out the optimum level of output, Q.

.....

.....

.....

.....

3. During the last period, the sum of average profit and fixed costs for a firm totaled ₹ 1, 00,000. Unit sales were 10,000. If variable cost per unit was ₹ 4, what was the selling price of a unit of output? How much would profit change if the firm produced and sold 11,000 units of output? (Assume average variable cost remains at ₹ 4 per unit).

.....

.....

.....

.....

8.11 SUMMARY

Cost concepts are important for decision-making but neither the accounting approach nor the economic approach is completely acceptable when decision-making is involved. Costs must be considered in various ways, depending on the decision at hand. Both traditional economists and traditional accountants have tended to be fairly dogmatic in their definitions of costs. On the other hand, managerial economists want a flexible approach. All the cost concepts need to be considered in such a way so as to help make sound decisions. The decision maker should try to discover the “relevant” costs by asking what costs are relevant to a particular decision at hand, and the decision maker is not necessarily bound by traditional concepts constructed for other purposes.

In short run, the total cost consists of fixed and variable costs. A firm’s marginal cost is the additional variable cost associated with each additional unit of output. The average variable cost is the total variable cost divided by the number of units of output. When there is a single variable input, the

presence of diminishing returns determines the shape of cost curves. In particular, there is an inverse relationship between the marginal product of the variable input and the marginal cost of production. The average variable cost and average total cost curves are U-shaped. The short run marginal cost curve increases beyond a certain point, and cuts both average total cost curve and average variable cost curve from below at their minimum points.

These cost concepts and analysis have a lot of applications in real world decision-making process such as optimum output, break even output, profit contribution, operating leverage, etc.

8.12 SELF ASSESSMENT QUESTIONS

1. Differentiate between accounting costs and economic costs?
2. Take a firm you are working with or know its nature. Make a list of relevant cost concepts from the standpoint of an (a) accountant and (b) economist.
3. What is the significance of opportunity cost in managerial decision-making?
4. What is short run cost analysis? For what type of decisions is it useful?
5. What are marginal costs and incremental costs? What is the difference between these two cost concepts?
6. A pharmaceutical company has spent ₹5 crores on developing and testing a new antibiotic drug. The head of the marketing department now estimates that it will cost ₹3 crores in advertising to launch this new product. Total revenue from all future sales is estimated at ₹6crores, and therefore, total costs will exceed revenue by ₹2 crores. He recommends that this product be dropped from the firm's product offerings. What is your reaction to this recommendation? The head of the accounting department now indicates that ₹3.5 crores of corporate overhead expenses also will be assigned to this product if it is marketed. Does this new information affect your decision? Explain.
7. When Mr. Kapoor's father gave him a new Truck costing ₹30 lakhs. Recently Mr. Kapoor was boasting to some of his friends that his revenues were typically ₹150,000 per month, while his operating costs (fuel, maintenance, and depreciation) amounted to only ₹1,20,000 per month. A truck identical to Mr. Kapoor's Truck is available on a monthly rent of ₹35000. If Mr. Kapoor was driving trucks for someone else, he would earn ₹5000 per month.
 - a. How much are Mr. Kapoor's explicit costs per month? How much are his implicit costs per month?
 - b. What is the amount of the opportunity cost of the resources used by Mr. Kapoor each month?

- c. Mr. Kapoor is proud of the fact that he is generating a net cash flow of ₹30000 ($= ₹1,50,000 - ₹1,20,000$) per month, since he would only be earning ₹5000 per month if he were working for some else. What advice would you give Mr. Kapoor?

9. The following table pertains to Savitha Company. Fill in the blanks below:

Output	Total Cost	Total Fixed Cost	Total Variable Cost	Average Total Cost	Average Fixed Cost	Average Variable Cost	Marginal Cost
100	260		60				
200							0.30
300						0.50	
400				1.05			
500			360				
600							3.00
700						1.60	
800	2040						

10. Suppose that a local metal fabricator has estimated its short run total cost function and total revenue function as

$$TC = 1600 + 100Q + 25Q^2$$

$$TR = 500Q$$

What is the breakeven amount of output? How might the company go about reducing the break even rate if it does not feel that it can sell the estimated amount in the market place?

11. A TV company sells color TV sets at ₹15,000 each. Its fixed costs are ₹30,000, and its average variable costs are ₹10,000 per unit. Draw its breakeven graph, and then determine its breakeven rate of production.
12. The Bright Electronics is producing small electronic calculators. It wants to determine how many calculators it must sell in order to earn a profit of ₹10,000 per month. The price of each calculator is ₹300, the fixed costs are ₹5,000 per month, and the average variable cost is ₹100.

- a. What is the required sales volume?

- b. If the firm were to sell each calculator at a price of ₹350 rather than ₹ 300, what would be the required sales volume?
- c. If the price is ₹ 350, and if average variable cost is ₹ 85 rather than ₹100, what would be the required sales volume?

8.13 FURTHERS READINGS

Adhikary, M. (1987). *Managerial Economics* (3rd ed.). Khosla Publishers, Delhi.

Maddala, G. S., & Miller, E. M. (1989). *Micro Economics: Theory and Applications*. McGraw-Hill, New York.

Mote, V. L., Paul, S., & Gupta, G. S. (2016). *Managerial Economics: Concepts and Cases*. Tata McGraw-Hill, New Delhi.

Mehta, P.L.(2016) *Managerial Economics: Analysis, Problems and Cases*. Sultan Chand & Sons, New Delhi.

Dholakia, R., & Oza, A. N. (1996). *Microeconomics for Management Students*. Oxford University Press, Delhi.

UNIT 9 LONG RUN COST ANALYSIS

Objectives

After studying this unit, you should be able to:

- analyze the behavior of costs in long run;
- comprehend the different sources of economies of scale;
- explain various functional forms of production and costs;
- understand empirical determination of these theoretical functions;
- identify managerial uses of such empirical estimates.

Structure

- 9.1 Introduction
- 9.2 Long-run Cost Functions
- 9.3 Economies and Diseconomies of Scale
- 9.4 Learning Curve
- 9.5 Economies of Scope
- 9.6 Cost Function and its Determinants
- 9.7 Estimation of Cost Function
- 9.8 Empirical Estimates of Cost Function
- 9.9 Managerial Uses of Cost Function
- 9.10 Summary
- 9.11 Self-Assessment Questions
- 9.12 Further Readings

9.1 INTRODUCTION

In unit 8, you have learnt about different cost concepts used by managers in decision-making process, the relationship between these concepts, and the distinction between accounting costs and economic costs, and short run cost analysis and its applications in managerial decision making. We will continue the analysis of costs in this unit also, long term cost analysis will be discussed.

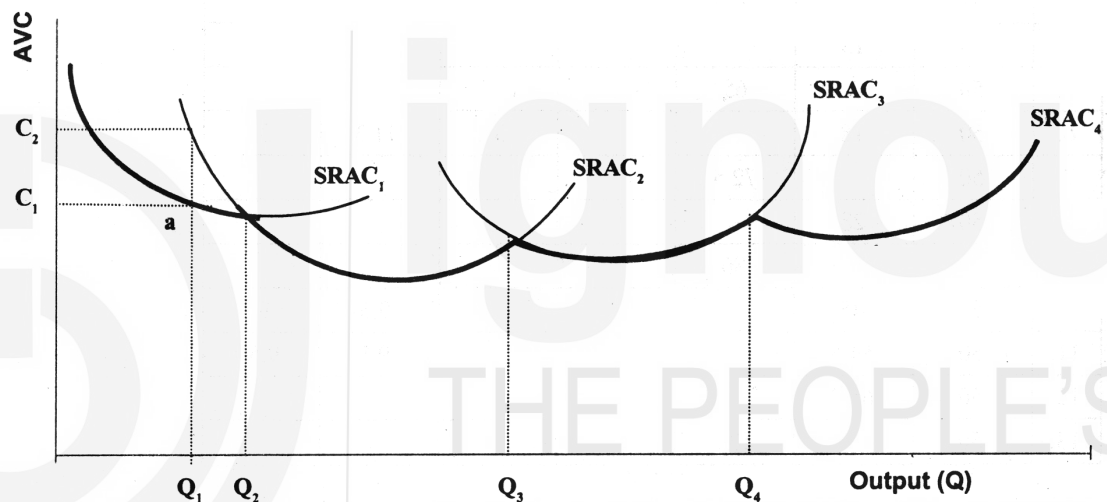
In the process of decision-making, a manager should understand clearly the relationship between the inputs and output on one hand and output and costs on the other. Estimation of cost curves will help production manager in understanding the nature and shape of cost curves and taking useful decisions. Both short run cost function and the long run cost function must be estimated, since both sets of information will be required for some vital decisions. Knowledge of the short run cost functions allows the decision makers to judge the optimality of present output levels and to solve decision problems of production manager. Knowledge of long run cost functions is important when considering the expansion or contraction of plant size, and

for confirming that the present plant size is optimal for the output level that is being produced. In the present Unit, we will discuss different approaches to examination of cost functions, analysis of some empirical estimates of these functions, and managerial uses of the estimated functions.

9.2 LONG-RUN COST FUNCTIONS

In the long run, all inputs are variable, and a firm can have a number of alternative plant sizes and levels of output that it wants. There are no fixed cost functions (total or average) in the long run, since no inputs are fixed. A useful way of looking at the long run is to consider it a planning horizon. The long run cost curve is also called *planning curve* because it helps the firm in future decision-making process.

Figure 9.1: Short-Run and Long Run Average Cost Curves



The long run cost output relationship can be shown with the help of a long run cost curve. The *long run average cost curve* (LRAC) is derived from *short run average cost curves* (SRAC). Let us illustrate this with the help of a simple example. A firm faces a choice of production with three different plant sizes viz. plant size-1 (small size), plant size-2 (medium size), plant size-3 (large size), and plant size-4 (very large size). The short run average cost functions shown in Figure 9.1 (SRAC₁, SRAC₂, SRAC₃, and SRAC₄) are associated with each of these plants discrete scale of operation. The long run average cost function for this firm is defined by the minimum average cost of each level of output. For example, output rate Q_1 could be produced by the plant size-1 at an average cost of C_1 or by plant size-2 at a cost of C_2 .

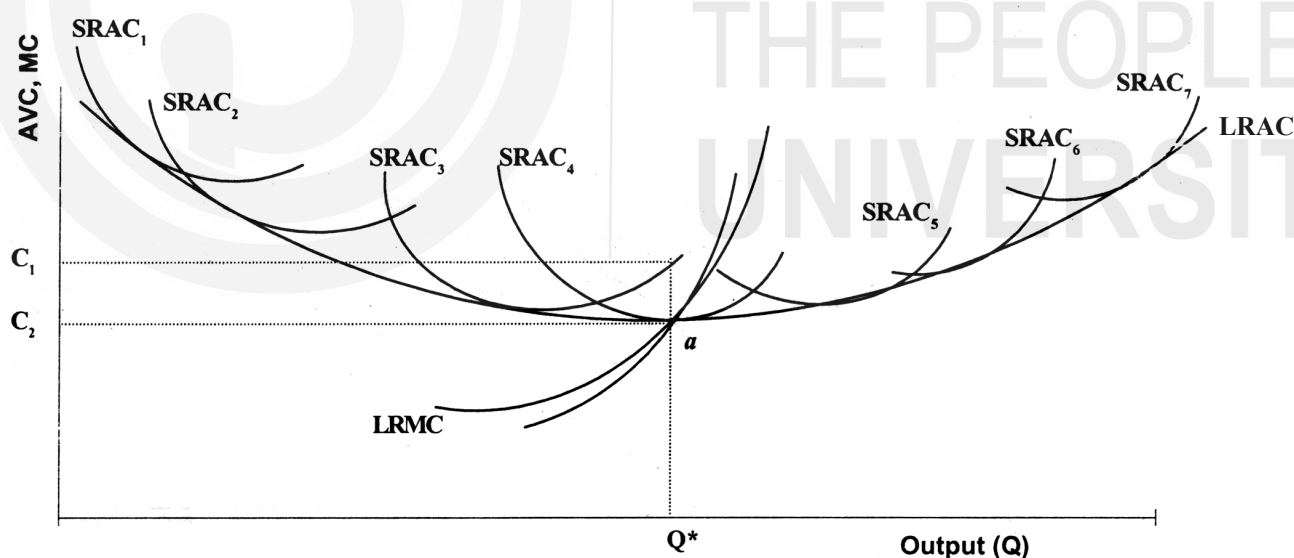
Clearly, the average cost is lower for plant size-1, and thus point a is one point on the long run average cost curve. By repeating this process for various rates of output, the long run average cost is determined. For output rates of zero to Q_2 plant size-1 is the most efficient and that part of SRAC₁ is part of the long run cost function. For output rates of Q_2 to Q_3 plant size-2 is the most efficient, and for output rates Q_3 to Q_4 , plant size-3 is the most

efficient. The scallop-shaped curve shown in bold face in Figure 9.1 is the long run average cost curve for this firm. This bold faced curve is called an **envelope curve** (as it envelopes short run average cost curves). Firms plan to be on this envelope curve in the long run. Consider a firm currently operating plant size-2 and producing Q_1 units at a cost of C_2 per unit. If output is expected to remain at Q_1 , the firm will plan to adjust to plant size-1, thus reducing average cost to C_1 .

Most firms will have many alternative plant sizes to choose from, and there is a short run average cost curve corresponding to each. A few of the short run average cost curves for these plants are shown in Figure 9.2, although many more may exist. Only one point of a very small arc of each short run cost curve will lie on the long run average cost function. Thus, long run average cost curve can be shown as the smooth U-shaped curve. Corresponding to this long run average cost curve is a long run marginal cost (LRMC) curve, which intersects LRAC at its minimum point a , which is also the minimum point of short run average cost curve 4 (SRAC₄). Thus, at a point a and only at a point a , the following unique result occurs:

$$\text{SRAC} = \text{SRMC} \quad \text{when} \quad \text{LRAC} = \text{LRMC}$$

Figure 9.2 : Short – Run and Long – Run Average Cost and Marginal Cost Curves



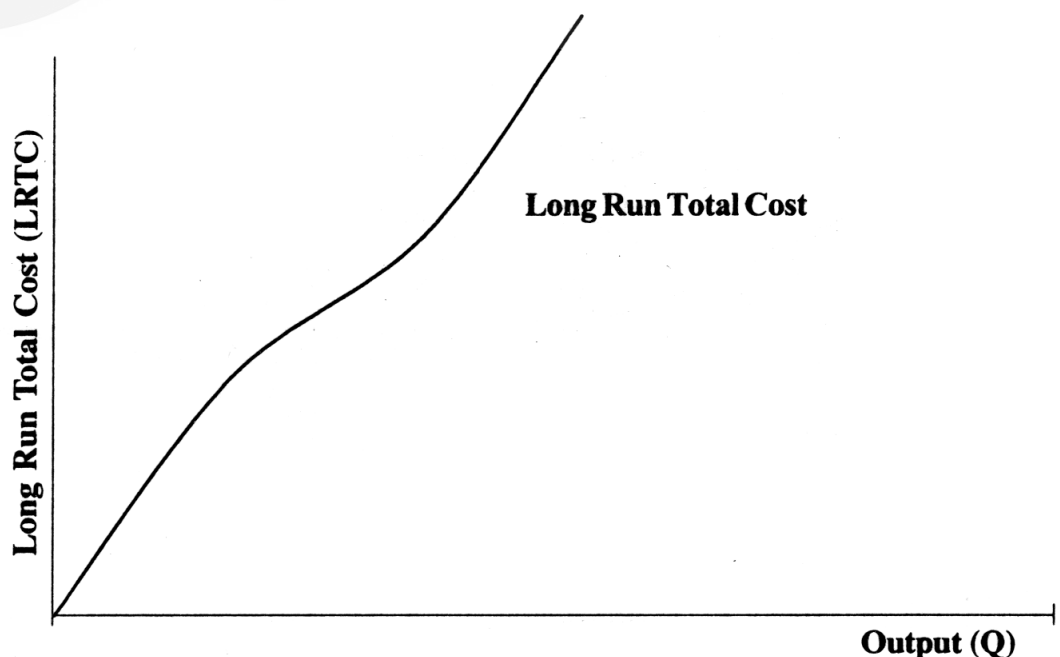
The long run cost curve serves as a long run planning mechanism for the firm. It shows the least per unit cost at any output can be produced after the firm has had time to make all appropriate adjustments in its plant size. For example, suppose that the firm is operating on short run average cost curve SRAC₃ as shown in Figure 9.2, and the firm is currently producing an output of Q^* . By using SRAC₃, it is seen that the firm's average cost is C_2 . Clearly, if projections of future demand indicate that the firm could expect to continue selling Q^* units per period at the market price, profit could be increased

significantly by increasing the scale of plant to the size associated with short run average cost curve $SRAC_4$. With this plant, average cost for an output rate of Q^* would be C_2 and the firm's profit per unit would increase by $C_2 - C_1$. Thus, total profit would increase by $(C_2 - C_1) \cdot Q^*$.

The U-shape of the LRAC curve reflects the laws of returns to scale. According to these laws, the cost per unit of production decreases as plant size increases due to the economies of scale, which the larger plant sizes make possible. But the economies of scale exist only up to a certain size of plant, known as the optimum plant size where all possible economies of scale are fully exploited. Beyond the optimum plant size, diseconomies of scale arise due to managerial inefficiencies. As plant size increases beyond a limit, the control, the feedback of information at different levels and decision-making process becomes less efficient. This makes the LRAC curve turn upwards. Given the LRAC in Figure 9.2, we can say that there are increasing returns to scale up to Q^* and decreasing returns to scale beyond Q^* . Therefore, the point Q^* is the point of optimum output and the corresponding plant size-4 is the optimum plant size. If you have long run average cost of producing a given output, you can readily derive the long run total cost (LRTC) of the output, since the long run total cost is simply the product of long run average cost and output. Thus, $LRTC = LRAC \cdot Q$.

Figure 9.3 shows the relationship between long run total cost and output. Given the long run total cost function you can readily derive the long run marginal cost function, which shows the relationship between output and the cost resulting from the production of the last unit of output, if the firm has time to make the optimal changes in the quantities of all inputs used.

Figure 9.3 : Long Run Total Cost Function



Activity 1

Long Run Cost Analysis

1. Explain why short run marginal cost is greater than long run marginal cost beyond the point at which they are equal?

.....

.....

.....

.....

.....

2. Explain why short run average cost can never be less than long run average cost?

.....

.....

.....

.....

.....

3. Why are all costs variable in the long run?

.....

.....

.....

.....

.....

4. Why is the long run average cost curve called an “envelope curve”? Why cannot the long run marginal cost curve be an envelope as well?

.....

.....

.....

.....

.....

5. What do you understand by “cost-efficiency”? Draw a long run cost diagram and explain.

.....

.....

.....

.....

.....

6. Economists frequently say that the firm plans in the long run and operates in the short run. Explain.

.....

.....

.....

.....

.....

.....

9.3 ECONOMIES AND DISECONOMIES OF SCALE

We have seen in the preceding section that larger plant will lead to lower average cost in the long run. However, beyond some point, successively larger plants will mean higher average costs. Exactly, why is the long run average cost (LRAC) curve U-shaped? What determines the shape of LRAC curve? This point needs further explanation.

It must be emphasized here that the law of diminishing returns is not applicable in the long run as all inputs are variable. Also, we assume that resource prices are constant. What then, is our explanation? The U-shaped LRAC curve is explainable in terms of what economists call *economies of scale* and *diseconomies of scale*.

Economies and diseconomies of scale are concerned with behaviour of average cost curve as the plant size is increased. If LRAC declines as output increases, then we say that the firm enjoys economies of scale. If, instead, the LRAC increases as output increases, then we have diseconomies of scale. Finally, if LRAC is constant as output increases, then we have constant returns to scale implying we have neither economies of scale nor diseconomies of scale. Economies of scale explain the down sloping part of the LRAC curve. As the size of the plant increases, LRAC typically declines over some range of output for a number of reasons. The most important is that, as the scale of output is expanded, there is greater potential for specialization of productive factors. This is most notable with regard to labour but may apply to other factors as well. Other factors contributing to declining LRAC include ability to use more advanced technologies and more efficient capital equipment; managerial specialization; opportunity to take advantage of lower costs (discounts) for some inputs by purchasing larger quantities; effective utilization of by products, etc.

But, after sometime, expansion of a firm's output may give rise to diseconomies, and therefore, higher average costs. Further expansion of output beyond a reasonable level may lead to problems of overcrowding of labour, managerial inefficiencies, etc., pushing up the average costs.

The economies of scale and diseconomies of scale are sometimes called as *internal economies* of scale and *internal diseconomies* of scale respectively. This is because the changes in long run average costs result solely from the individual firm's adjustment of its output. On the other hand, there may exist external economies of scale. The *external economies* also help in cutting down production costs. With the expansion of an industry, certain specialized firms also come up for working up the by-products and waste materials. Similarly, with the expansion of the industry, certain specialized units may come up for supplying raw material, tools, etc., to the firms in the industry. Moreover, they can combine together to undertake research etc., whose benefit will accrue to all firms in the industry. Thus, a firm benefits from expansion of the industry as a whole. These benefits are external to the firm, in the sense that these have arisen not because of any effort on the part of the firm but have accrued to it due to expansion of industry as a whole. All these external economies help in reducing production costs.

Economies of scale are often measured in terms of cost- output elasticity, E_c .

E_c is the percentage change in the average cost of production resulting from a one percent increase in output:

$$E_c = (\Delta TC/TC) / (\Delta Q/Q) = (\Delta TC/\Delta Q) / (TC/Q) = MC/AC$$

Clearly, E_c is equal to one when marginal and average costs are equal. This means costs increase proportionately with output, and there are neither economies nor diseconomies of scale. When there are economies of scale MC will be less than AC (both are declining) and E_c is less than one. Finally, when there are diseconomies of scale, MC is greater than AC, and E_c is greater than one.

ECONOMIES OF SCALE

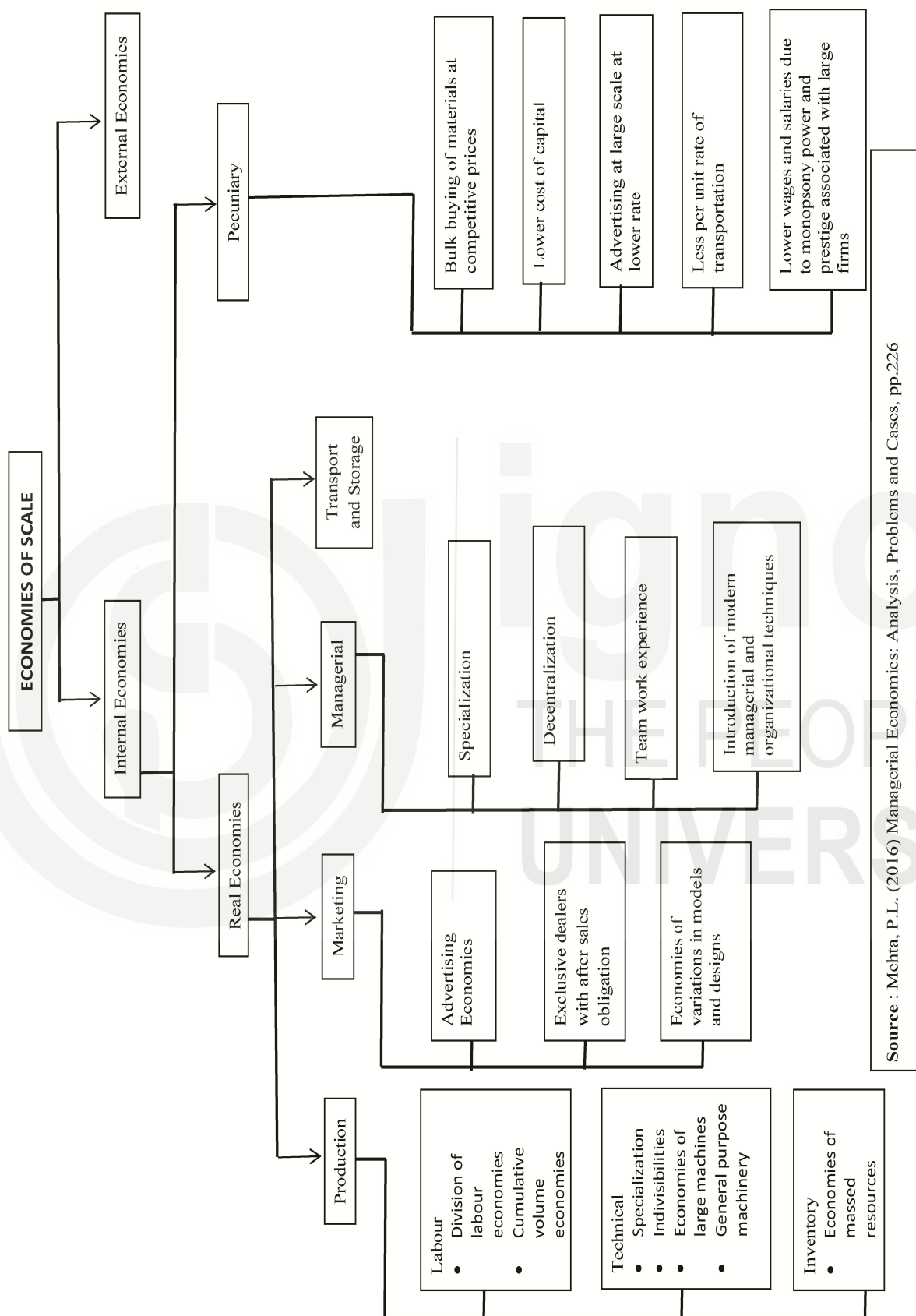
Internal economies

Real and Pecuniary are two types of internal economies. When quantity of inputs used decrease for a particular level of output, it gives rise to **real economies**. When large quantities of input are bought and large quantities of output are sold, there are savings in the cost of inputs and distribution costs due to bulk buying and selling, these savings are known as **pecuniary economies**.

A) REAL ECONOMIES OF SCALE

1. Production economies

Production economies arise from different factors of production such as :-



Source : Mehta, P.L. (2016) Managerial Economics: Analysis, Problems and Cases, pp.226

- **Labour Economies**

Large scale of production leads to division of labour and specialization, leading to reduction in costs and time. Due to large scale production, technical personnel acquire significant experience and this 'cumulative volume' has positive effect on production and costs, leading to high productivity and lower costs for large level of outputs.

- **Technical Economies**

Technical economies are related to fixed capital which includes machinery and equipments. Specialization and indivisibility of machines and equipments leads to reduction in per unit costs. Machines have a property of indivisibility so, when large production (maximum capacity of machine) is achieved, machine cost is divided between large units.

In large size firms, there are large machines which perform different continuous automatic processes which saves time and cost of labour as different machines for different steps are not required. And with increase in scale, general purpose machinery is left to set up more which reduces its set up costs.

- **Inventory economies**

The main aim of inventories is to meet random fluctuations in supply and demand of inputs & output respectively. When size of firm increases, these fluctuations are eased out due to large quantity of resources.

2) Marketing economies

With increase in output, advertising, R&D (such as developing new models and designs) expenditures etc. are spread over and expenditure per unit decreases considerably.

3) Managerial economies

Managerial economies arise due to various reasons, most important of them are specialization, decentralization, team work experience and modern managerial techniques. When scale of operations rises, division of managerial tasks (production, sales, finance etc.) is done and this leads to specialization as same task is performed again and again. Team work experience helps in improving decision making ability of managers and with decentralization decision making becomes faster leading to managerial efficiency.

4) Transport and storage economies

Storage costs decrease with increase in the size of output. But transportation costs keep falling for some extent and then they become constant, making a L shaped curve.

B) PECUNIARY ECONOMIES OF SCALE

These economies include all the savings and discounts obtained by firm because of its large size. Few examples of these are:

- Raw material prices become competitive when there is bulk buying.

- Banks offer lower rates of interest to large firms, ultimately it comes down to lower cost of capital.
- Large firms have need of large-scale advertising, they are offered better prices.
- Transportation costs turn out lower for bulk transportation.
- Labour costs can be saved by large firms, if they attain a position, which gives them monopsony power due to large size or prestige in market.

EXTERNAL ECONOMIES

External economies are those which are not particular to one particular firm rather these are shared by all firms operating in a particular industry. These are external to the firm as they do not arise due to efforts of any one particular firm, they arise when whole industry expands. These can be economies of information, economies of concentration, economies of disintegration etc. These external economies also help in reduction of production costs just like internal economies.

DISECONOMIES OF SCALE

Diseconomies of scale occur, at a level of output where one unit increase in output leads to increase in marginal costs. At this level, diminishing returns are there and economies of scale are ruled out by diseconomies of scale. Diseconomies of scale generally arise due to **human and behavioural problems** instead of technical factors.

There are several reasons for diseconomies of scale, some of them have been listed below:

Morale and motivation problems

When the size of firm increases more and more employees join the organization which reduces personal connection with owners and management which leads to reduction in motivation among employees. Reduction in motivation leads to reduction in productivity as well.

Communication problems

Effective and properly defined communication channels are very important in any business organization. When a firm increases its size, there are increased layers of command, it may create problems in communication and a message can be distorted or lost. So, the formal communication of objectives and goals will also suffer in this scenario which will lead to reduction in productivity.

In short, it can be concluded that diseconomies of scale will come into picture when with increase in size, management faces tough challenges in regards to communication, motivation & morale of employees etc. and they are not able to overcome those challenges.

1. Distinguish between internal and external economies of scale. Give examples.

.....

.....

.....

.....

.....

2. Distinguish between real and pecuniary economies of scale.

.....

.....

.....

.....

.....

3. Explain the reasons of diseconomies of scale with examples.

.....

.....

.....

.....

.....

9.4 LEARNING CURVE

Learning curve denotes the relationship between labour cost and additional units of output. It basically revolves around the concept that, with practice workers improve and ultimately, cost per unit of producing one more units of output decreases. **Learning curve effect** is this reduction in cost due to learning process of employees over a period of time.

9.5 ECONOMIES OF SCOPE

According to the concept of economies of scale, cost advantages follow the increase in volume of production or what is called the scale of output. On the other hand, according to the concept of economies of scope, such cost advantages may follow from a variety of output. For example, many firms produce more than one product and the products are closely related to one another — an automobile company produces scooters and cars, and a university produces teaching and research. A firm is likely to enjoy production or cost advantages when it produces two or more products. These advantages could result from the joint use of inputs or production facilities, joint marketing programs, or possibly the cost savings of a common

administration. Examples of joint products are mutton and wool, eggs and chicken, fertilizer, etc.

Therefore, economies of scope exist when the cost of producing two (or more) products jointly is less than the cost of producing a single product. To measure the degree to which there are economies of scope, we should know what percentage of the cost of production is saved when two (or more) products are produced jointly rather than individually. The following equation gives the degree of economies of scope (SC) that measures the savings in cost:

$$SC = \frac{C(Q_1) + C(Q_2) - C(Q_1 + Q_2)}{C(Q_1 + Q_2)}$$

Here, $C(Q_1)$ represents the cost of producing output Q_1 , $C(Q_2)$ the cost of producing output Q_2 , and $C(Q_1, Q_2)$ the joint cost of producing both outputs ($Q_1 + Q_2$).

For example, a firm produces 10000 TV sets and 5000 Radio sets per year at a cost of Rs.8.40 crores, and another firm produces 10000 TV sets only, then the cost would be Rs.10.00 crores, and if it produced 5000 Radio sets only, then the cost would be Rs. 0.50 crores. In this case, the cost of producing both the TV and Radio sets is less than the total cost of producing each separately. Thus, there are economies of scope. Thus,

$$SC = \frac{10.00 + 0.50 - 8.40}{8.40} = 0.25$$

Which means that there is a 25% saving of cost by going for joint production.

With economies of scope, the joint cost is less than the sum of the individual costs, so that SC is greater than 0. With diseconomies of scope, SC is negative. In general, the larger the value of SC, the greater is the economies of scope.

Activity 3

1. Distinguish between economies of scale and economies of scope using examples.

.....

.....

.....

.....

.....

9.6 COST FUNCTION AND ITS DETERMINANTS

Cost function expresses the relationship between cost and its determinants such as the size of plant, level of output, input prices, technology, managerial efficiency, etc. In a mathematical form, it can be expressed as,

$$C = f(S, O, P, T, E, \dots)$$

Where, C = cost (it can be unit cost or total cost)

S = plant size

O = output level

P = prices of inputs used in production

T = nature of technology

E = managerial efficiency

Determinants of Cost Function

The cost of production depends on many factors and these factors vary from one firm to another firm in the same industry or from one industry to another industry. The main determinants of a cost function are:

- a) plant size
- b) output level
- c) prices of inputs used in production,
- d) nature of technology
- e) managerial efficiency

We will discuss briefly the influence of each of these factors on cost.

- a) **Plant size:** Plant size is an important variable in determining cost. The scale of operations or plant size and the unit cost are inversely related in the sense that as the former increases, unit cost decreases, and vice versa. Such a relationship gives downward slope of cost function depending upon the different sizes of plants taken into account. Such a cost function gives primarily engineering estimates of cost.
- b) **Output level:** Output level and total cost are positively related, as the total cost increases with increase in output and total cost decreases with decrease in output. This is because increased production requires increased use of raw materials, labour, etc., and if the increase is substantial, even fixed inputs like plant and equipment, and managerial staff may have to be increased.
- c) **Price of inputs:** Changes in input prices also influence cost, depending on the relative usage of the inputs and relative changes in their prices. This is because more money will have to be paid to those inputs whose prices have increased and there will be no simultaneous reduction in the costs from any other source. Therefore, the cost of production varies directly with the prices of production.

- d) **Technology:** Technology is a significant factor in determining cost. By definition, improvement in technology increases production leading to increase in productivity and decrease in production cost. Therefore, cost varies inversely with technological progress. Technology is often quantified as capital-output ratio. Improved technology is generally found to have higher capital-output ratio.
- e) **Managerial efficiency:** This is another factor influencing the cost of production. More the managerial efficiency less the cost of production. It is difficult to measure managerial efficiency quantitatively. However, a change in cost at two points of time may explain how organizational or managerial changes within the firm have brought about cost efficiency, provided it is possible to exclude the effect of other factors

9.7 ESTIMATION OF COST FUNCTION

Several methods exist for the measurement of the actual cost-output relation for a particular firm or a group of firms, but the three broad approaches - accounting, engineering and econometric – are the most important and commonly used.

Accounting Method

This method is used by the cost accountants. In this method, the cost-output relationship is estimated by classifying the total cost into fixed, variable and semi-variable costs. These components are then estimated separately. The average variable cost, the semi-variable cost which is fixed over a certain range of output, and fixed costs are determined on the basis of inspection and experience. The total cost, the average cost and the marginal cost for each level of output can then be obtained through a simple arithmetic procedure.

Although, the accounting method appears to be quite simple, it is a bit cumbersome as one has to maintain a detailed breakdown of costs over a period to arrive at good estimates of actual cost-output relationship. One must have experience with a wide range of fluctuations in output rate to come up with accurate estimates.

Engineering Method

The engineering method of cost estimation is based directly on the physical relationship of inputs to output, and uses the price of inputs to determine costs. This method of estimating real world cost function rests clearly on the knowledge that the shape of any cost function is dependent on: (a) the production function and (b) the price of inputs.

We have seen earlier in Unit – 7 while discussing the estimation of production function that for a given the production function and input prices,

the optimum input combination for a given output level can be determined. The resultant cost curve can then be formulated by multiplying each input in the least cost combination by its price, to develop the cost function. This method is called engineering method as the estimates of least cost combinations are provided by engineers.

The assumption made while using this method is that both the technology and factor prices are constant. This method may not always give the correct estimate of costs as the technology and factor prices do change substantially over a period of time. Therefore, this method is more relevant for the short run. Also, this method may be useful if good historical data is difficult to obtain. But this method requires a sound understanding of engineering and a detailed sampling of the different processes under controlled conditions, which may not always be possible.

Econometric Method

This method is also sometimes called statistical method and is widely used for estimating cost functions. Under this method, the historical data on cost and output are used to estimate the cost-output relationship. The basic technique of regression is used for this purpose. The data could be a time series data of a firm in the industry or of all firms in the industry or a cross-section data for a particular year from various firms in the industry.

Depending on the kind of data used, we can estimate short run or long run cost functions. For instance, if time series data of a firm whose output capacity has not changed much during the sample period is used, the cost function will be short run. On the other hand, if cross-section data of many firms with varying sizes, or the time series data of the industry as a whole is used, the estimated cost function will be the long run one.

The procedure for estimation of cost function involves three steps. First, the determinants of cost are identified. Second, the functional form of the cost function is specified. Third, the functional form is chosen and then the basic technique of regression is applied to estimate the chosen functional form.

Functional Forms of Cost Function

The following are the three common functional forms of cost function in terms of total cost function (TC).

Linear cost function: $TC = a_1 + b_1Q$

Quadratic cost function: $TC = a_2 + b_2Q + c_2Q^2$

Cubic cost function: $TC = a_3 + b_3Q + c_3Q^2 + d_3Q^3$

where, $a_1, a_2, a_3, b_1, b_2, b_3, c_2, c_3, d_3$ are constants.

When all the determinants of cost are chosen and the data collection is complete, the alternative functional forms can be estimated by using

regression software package on a computer. The most appropriate form of the cost function for decision-making is then chosen on the basis of the principles of economic theory and statistical inference.

Once the constants in the total cost function are estimated using regression technique, the average cost (AC) and marginal cost (MC) functions for chosen forms of cost function will be calculated. The TC, AC and MC cost functions for different functional forms of total cost function and their typical graphical presentation and interpretation are explained below.

A. Linear cost function

$$TC = a_1 + b_1Q$$

$$AC = (TC)/Q = (a_1/Q) + b_1$$

$$MC = \frac{d(TC)}{dQ} = b_1$$

The typical TC, AC, and MC curves that are based on a linear cost function are shown in Figure 9.4. These cost functions have the following properties: TC is a linear function, where AC declines initially and then becomes quite flat approaching the value of MC as output increases and MC is constant at b_1 .

B. Quadratic cost function

$$TC = a_2 + b_2Q + c_2Q^2$$

$$AC = (TC)/Q = (a_2/Q) + b_2 + c_2Q$$

$$MC = \frac{d(TC)}{dQ} = b_2 + 2c_2Q$$

The typical TC, AC, and MC curves that are based on a quadratic cost function are shown in Figure 9.5. These cost functions have the following properties: TC increases at an increasing rate; MC is a linearly increasing function of output; and AC is a U-shaped curve.

C. Cubic cost function

$$TC = a_3 + b_3Q + c_3Q^2 + d_3Q^3$$

$$AC = (TC)/Q = (a_3/Q) + b_3 + c_3Q + d_3Q^2$$

$$MC = \frac{d(TC)}{dQ} = b_3 + 2c_3Q + 3d_3Q^2$$

The typical TC, AC, and MC curves that are based on a cubic cost function are shown in Figure 9.6. These cost functions have the following properties: TC first increases at a decreasing rate up to output rate Q_1 in the Figure 9.6

and then increases at an increasing rate; and both AC and MC cost functions are U shaped functions.

The linear total cost function would give a constant marginal cost and a monotonically falling average cost curve. The quadratic function could yield a U-shaped average cost curve but it would imply a monotonically rising marginal cost curve. The cubic cost function is consistent both with a U-shaped average cost curve and a U-shaped marginal cost curve. Thus, to check the validity of the theoretical cost-output relationship, one should hypothesize a cubic cost function.

Figure 9.4: Cost Curves Based on Linear Cost Function

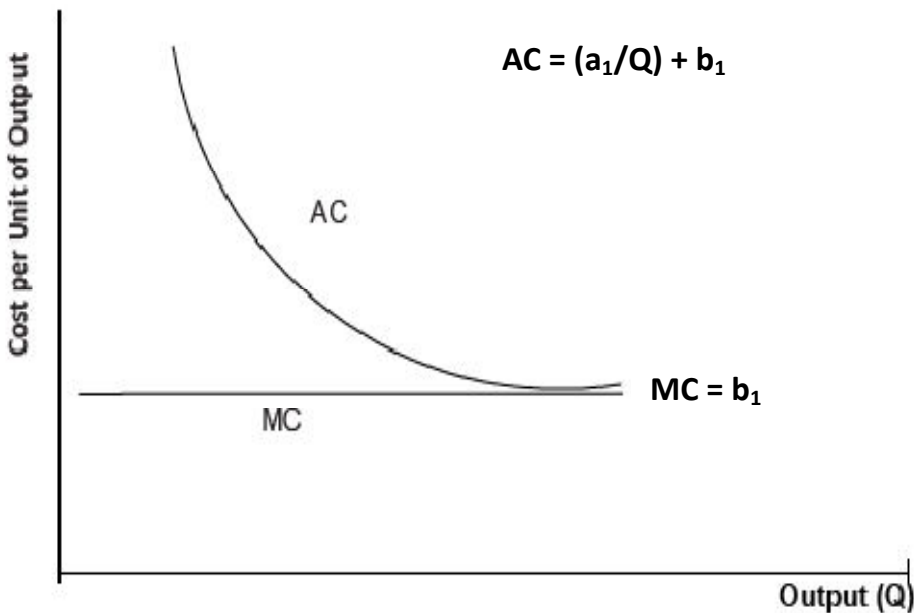
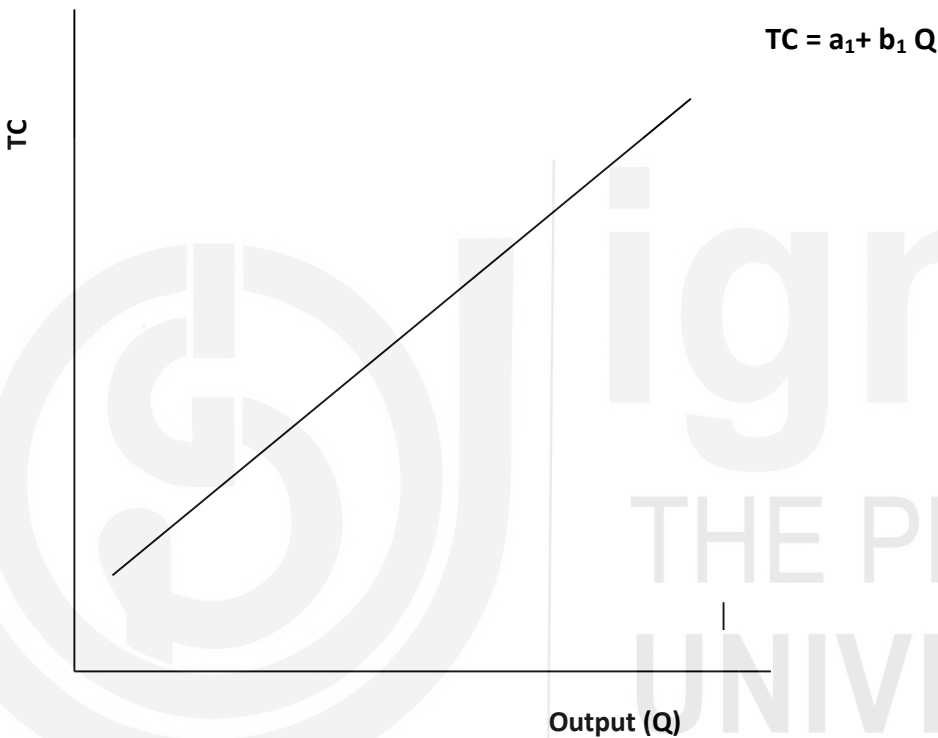


Figure 9.5: Cost Curves Based on Quadratic Cost Function

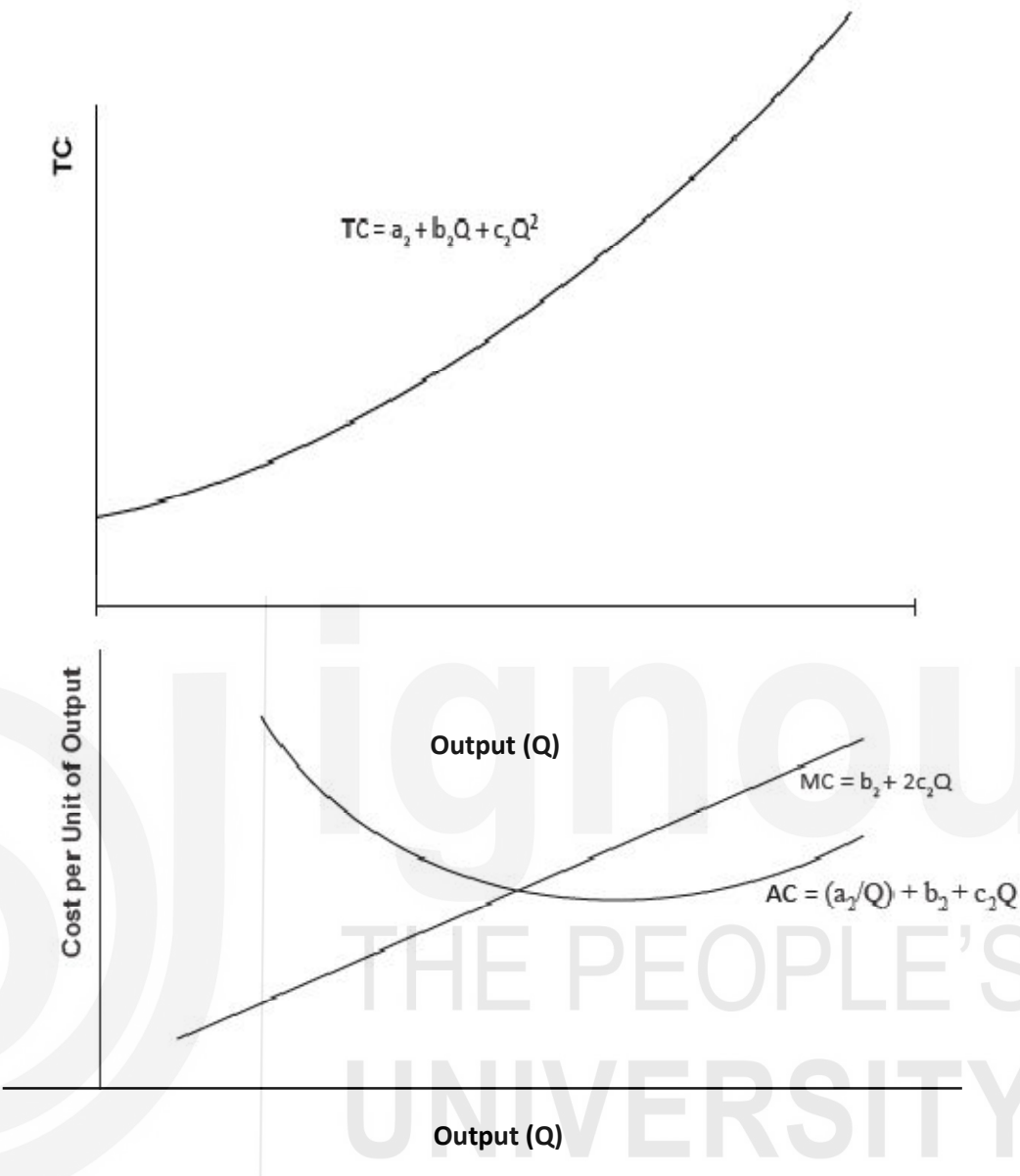
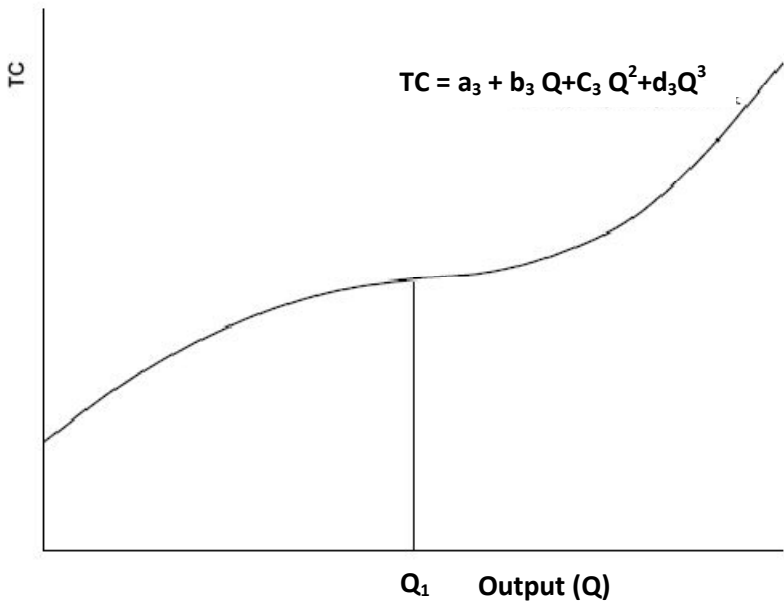
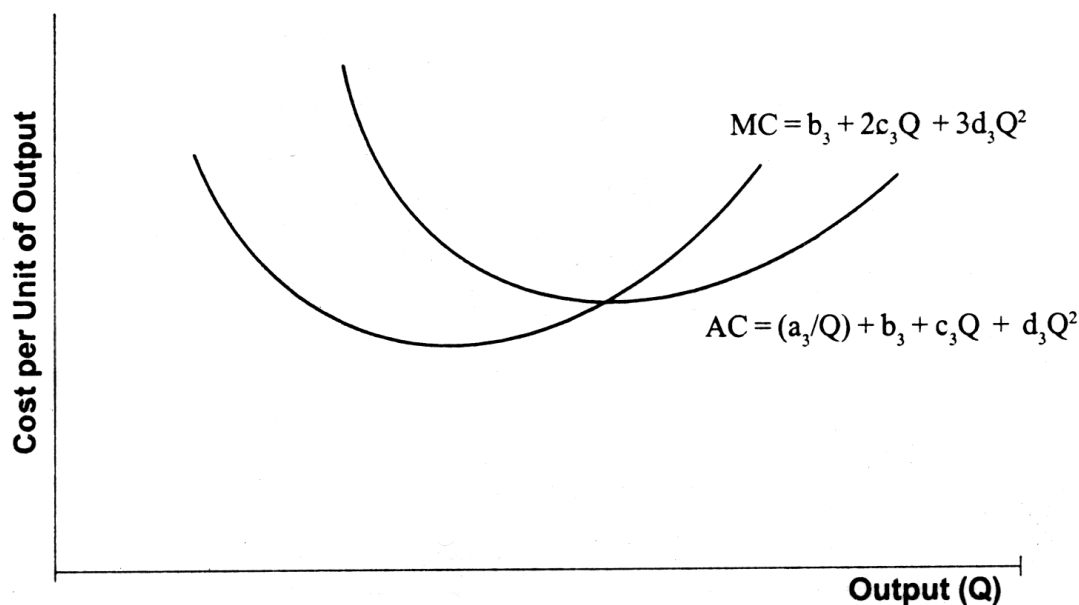


Figure 9.6 : Cost curves Based on Cubic Cost Function





An example of using estimated cost function:

Using the output-cost data of a chemical firm, the following total cost function was estimated using quadratic function:

$$TC = 1016 - 3.36Q + 0.021Q^2$$

- Determine average and marginal cost functions.
- Determine the output rate that will minimize average cost and the per unit cost at that rate of output.
- The firm proposed a new plant to produce nitrogen. The current market price of this fertilizer is Rs 5.50 per unit of output and is expected to remain at that level for the foreseeable future. Should the plant be built?

- The average cost function is

$$AC = (TC/Q) = (a_2/Q) + b_2 + c_2Q = (1016/Q) - 3.36 + 0.021Q$$

and the marginal cost function is

$$MC = \frac{d(TC)}{dQ} = b_2 + 2c_2Q = -3.36 + 2(0.021)Q = -3.36 + 0.042Q$$

- The output rate that results in minimum per unit cost is found by taking the first derivative of the average cost function, setting it equal to zero, and solving for Q.

$$\frac{d(AC)}{dQ} = \frac{a_2}{Q^2} + c_2 = -\frac{1016}{Q^2} + 0.021 = 0$$

$$\frac{1016}{Q^2} = 0.021; 0.021Q^2 = 1016; Q^2 = 48381; Q = 220$$

To find the cost at this rate of output, substitute 220 for Q in AC equation and solve it.

$$AC = (1016/Q) - 3.36 + 0.021Q = (1016/220) - 3.36 + (0.021 * 220)$$

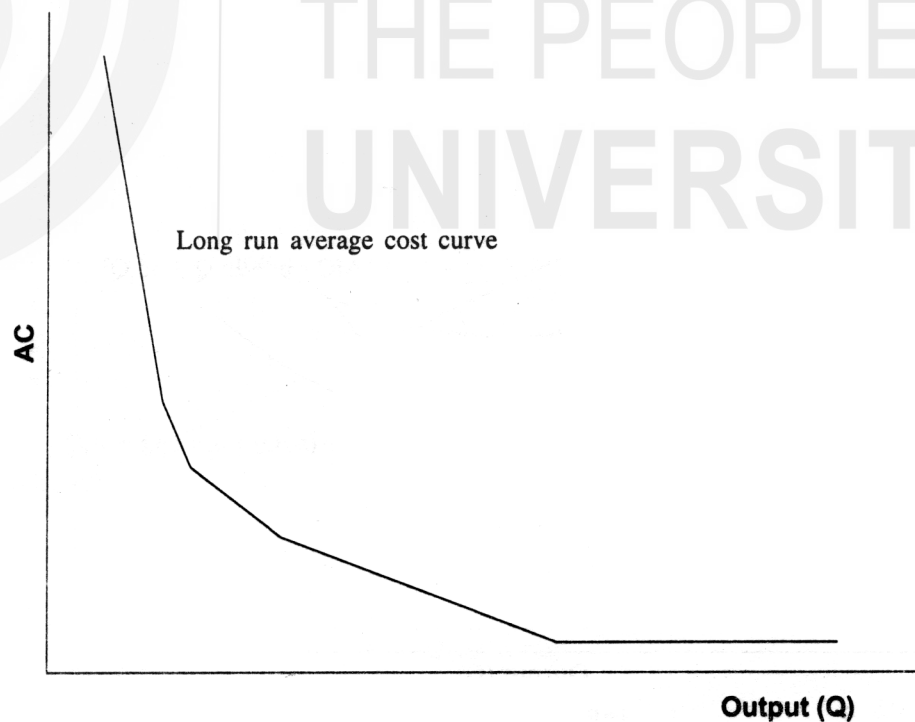
Rs. 5.88 per unit of output.

iii) Because the lowest possible cost is Rs. 5.88 per unit, which is Rs. 0.38 above the market price (Rs.5.50), the plant should not be constructed.

Short Run and Long Run Cost Function Estimation

The same sorts of regression techniques can be used to estimate short run cost functions and long run cost functions. However, it is very difficult to find cases where the scale of a firm has changed but technology and other relevant factors have remained constant. Thus, it is hard to use time series data to estimate long run cost functions. Generally, regression analysis based on cross section data has been used instead. Specially, a sample of firms of various sizes is chosen, and a firm's TC is regressed on its output, as well as other independent variables, such as regional differences in wage rates or other input prices.

Figure 9.7: Typical Long Run Average Cost Curve



Many studies of long run cost functions that have been carried out found that there are very significant economies of scale at low output levels, but that these economies of scale tend to diminish as output increases, and that the long run average cost function eventually becomes close to horizontal axis at

high output levels. Therefore, in contrast to the U-shaped curve in Figure 9.1, which is often postulated in micro economic theory, the long run average cost curve tends to be L-shaped, as shown in Figure 9.7.

Problems in Estimation of Cost Function

We confront certain problems while attempting to derive empirical cost functions from economic data. Some of these problems are briefly discussed below.

1. In collecting cost and output data we must be certain that they are properly paired. That is, the cost data applicable to the corresponding data on output.
2. We must also try to obtain data on cost and output during a time period when the output has been produced at relatively even rate. If for example, a month is chosen as the relevant time period over which the variables are measured, it would not be desirable to have wide weekly fluctuations in the rate of output. The monthly data in such a case would represent an average output rate that could disguise the true cost- output relationship. Not only should the output rate be uniform, but it also should be a rate to which the firm is fully adjusted. Furthermore, there should be no disruptions in the output due to external factors such as power failures, delays in receiving necessary supplies, etc. To generate the data necessary for a meaningful statistical analysis, the observations must include a wide range of rates of output. Observing cost-output data for the last 24 months, when the rate of output was the same each month, would provide little information concerning the appropriate cost function.
3. The cost data is normally collected and recorded by accountants for their own purposes and in a manner that it makes the information less than perfect from the perspective of economic analysis. While collecting historical data on cost, care must be taken to ensure that all explicit as well as implicit costs have been properly taken into account, and that all the costs are properly identified by time period in which they were incurred.
4. For situations in which more than one product is being produced with given productive factors, it may not be possible to separate costs according to output in a meaningful way. One simple approach of allocating costs among various products is based on the relative proportion of each product in the total output. However, this may not always accurately reflect the cost appropriate to each output.

5. Since prices change over time, any money value cost would therefore relate partly to output changes and partly to price changes. In order to estimate the cost-output relationship, the impact of price change on cost needs to be eliminated by deflating the cost data by price indices. Wages and equipment price indices are readily available and frequently used to 'deflate' the money cost.
6. Finally, there is a problem of choosing the functional form of equation or curve that would fit the data best. The usefulness of any cost function for practical application depends, to a large extent, on appropriateness of the functional form chosen. There are three functional forms of cost functions, which are popular, viz., linear, quadratic and cubic. The choice of a particular function depends upon the correspondence of the economic properties of the data to the mathematical properties of the alternative hypotheses of total cost function.

The accounting and engineering methods are more appropriate than the econometric method for estimating the cost function at the firm level, while the econometric method is more suitable for estimating the cost function at the industry or national level. There has been a growing application of the econometric method at the macro level and there are good prospects for its use even at the micro level. However, it must be understood that the three approaches discussed above are not competitive, but are rather complementary to each other. They supplement each other. The choice of a method therefore depends upon the purpose of study, time and expense considerations.

9.8 EMPIRICAL ESTIMATES OF COST FUNCTION

A number of studies using time series and cross-section data have been conducted to estimate short run and long run cost behavior of various industries. Table 9.1 lists a number of well-known studies estimating short run average and marginal cost curves. These and many other studies point one conclusion: in the short run a linear total variable cost function with constant marginal cost is the relationship that appears to describe best the actual cost conditions over the "normal" range of production. U-shaped average cost (AC) and marginal cost (MC) curves have been found, but are less prevalent than one might expect.

Table 9.1: Number of well-known studies estimating short run average and marginal cost curves

Name	Type of Industry	Findings
Dean (1936)	Furniture	Constant MC which failed to rise
Dean (1941)	Leather belts	No significant increases in MC
Dean (1941)	Hosiery	Constant MC which failed to rise
Dean (1942)	Department store	Declining or constant MC, depending on the department within the store
Ezekiel and Wylie (1941)	Steel	Declining MC but large variation
Hall and Hitch (1939)	Manufacturing	Majority have decreasing MC
Johnston (1960)	Electricity, multi product function food processing	“Direct” cost is a linear function of output, and MC is constant
Johnston (1960)	Electricity	Average total cost falls, then flattens, tending toward constant MC up to capacity
Mansfield and Wein (1958)	Railways	Constant MC
Yntema (1940)	Steel	Constant MC

Source: A.A. Walters, “Production and Cost Functions: An Econometric Survey”, *Econometrica*, January-February 1963, PP.49-54

Table 9.2 lists a number of well known, long run average cost studies. In some industries, such as light manufacturing (of baking products), economies of size are relatively unimportant and diseconomies set in rather quickly, implying that a small plant has cost advantages over a large plant. In other industries, such as meat packing or the production of household appliances, the long run average cost curve is found to be flat over an extended range of output, thereby indicating that a variety of different plant sizes are all more or less equally efficient. In some other industries such as electricity or metal (aluminum and steel) production, substantial economies of size are found, thereby implying that a large plant is most efficient. Rarely are substantial diseconomies of size found in empirical studies, perhaps because of firms recognizing that production beyond a certain range leads to sharply rising costs. Therefore, they avoid such situations if all possible by building additional plants.

Table 9.2: Number of well known, long run average cost studies

Name	Type of Industry	Findings
Alpert (1959)	Metal	Economies of scale up to some level of output per month; constant returns to scale and horizontal LRAC thereafter
Bain (1956)	Manufacturing	Small economies of scale for multi-plant firms
Gribbin (1953)	Gas (Great Britain)	LRAC of production declines as output rises
Holton (1956)	Retailing	LRAC L-shaped
Johnston (1960)	Life Assurance	LRAC declines
Johnston (1960)	Road passenger transport (Great Britain)	LRAC either falling or constant
Johnston (1960)	Electricity (Great Britain)	LRAC of production declines as output rises
Lomax (1951)	Gas (Great Britain)	LRAC of production declines as output rises
Lomax (1952)	Electricity (Great Britain)	LRAC of production declines as output rises
Moore (1959)	Manufacturing	Economies of scale prevail quite generally
Nerlove (1961)	Electricity (U.S.)	LRAC (excluding transmission costs) declines and then shows signs of increasing
Gupta (1968) *	Manufacturing (India)	L-shaped in 18 industries, U-shaped in 5 industries, and linear in 6 industries

Source: A.A. Walters, "Production and Cost Functions: An Econometric Survey", *Econometrica*, January-February 1963, PP.49-54.

*Vinod K. Gupta, "Cost Functions Concentration and Barriers to entry in twenty-nine Manufacturing Industries of India." *Journal of Industrial Economics*, November 1, 1968, 59-60

Activity 4

1. Pradeep Company's total variable function is as follows: $TVC = 50Q - 10Q^2 + Q^3$

Where Q is the number of units of output produced?

- a) What is the output level where marginal cost is a minimum?
- b) What is the output level where average variable cost is a minimum?
- c) What is the value of average variable cost and marginal cost at the output specified in the answer to part (b)?

.....

.....

.....

.....

.....

2. How would you reconcile the findings of Yntena with those of Ezekiel and Wylie?

.....

.....

.....

.....

.....

3. How would you explain the findings of Johnston (Electricity) in short run and long run?

.....

.....

.....

.....

.....

4. Production is related to costs. In fact, cost function can be derived from estimated production function. In view of empirical determination of production function, can you think of some limitations of statistical analysis relating to cost function?

.....

.....

.....

.....

.....

5. Despite the above limitations listed by you, an estimated cost function is useful to a manager. Can you think of some points to support this contention?

6. Some empirical studies have suggested that the marginal cost function is approximately horizontal, but conventional cost theory suggests that the marginal cost curve is U-shaped. Provide an explanation for this apparent inconsistency.

7. The ABC Manufacturing Company's short-run average cost function in the year 2000 is $AC=3+4Q$

Where AC is the firm's average cost (in Rs. per unit of the product), and Q is the output rate.

- Obtain the firm's short-run total cost function.
- Does the firm have any fixed costs? Explain.
- If the price of the firm's product is Rs. 3 per unit, is the firm making profits or losses? Explain.
- Derive the firm's marginal cost function.

9.9 MANAGERIAL USES OF COST FUNCTION

The estimated cost function can help managers to take meaningful decisions with regard to:

- determination of optimum plant size,
- determination of optimum output for a given plant, and

3. determination of a firm's supply curve.

The optimum plant size, as discussed earlier, is defined in terms of minimum costs per unit of output. In other words, an optimum plant is given by that value of K (plant size) for which the average cost is minimum. If the long run total cost curve is a cubic function, the resultant long run average cost curve will be a conventional U-shaped curve. The plant level at which the long run average cost is minimum will be of optimum size.

For a given plant, the optimum output level will be achieved at a point where the average cost is the least. This condition can be easily verified from the short run total cost function.

The level of output that a firm would like to supply to the market will depend on the price that can charge for its product. In other words, a firm's supply is a positive function of the product price. To get the firm's supply schedule, one needs to know the firm's cost function and its objectives.

Activity 5

1. Can you list some more managerial uses of cost function other than given in section 9.9?

.....

.....

.....

.....

.....

9.10 SUMMARY

In the long run, all inputs to the production process are variable. Thus, in the long run, total costs are identical to variable costs. The long run average cost function shows the minimum cost for each output level when a desired scale of plant can be built. The long run average cost curve is important to managers because it shows the extent to which larger plants have cost advantages over smaller ones. Economies or diseconomies of scale arise either due to the internal factors pertaining to the expansion of output by a firm, or due to the external factors such as industry expansion. In contrast, economies of scope result from product diversification. Thus, the scale-economies have reference to an increase in volume of production, whereas the scope-economies have reference to an improvement in the variety of products from the existing plant and equipment.

We have also discussed three forms of cost functions viz. linear cost function, quadratic cost function, and cubic cost function and their empirical estimates. Though, empirical estimates of both production functions and cost functions

have a lot of use for managerial decision making. There are conceptual and statistical problems in estimating such functions. But we understand that it will be sufficient for the manager if he knows how to interpret the estimates based on empirical research in her/his decision-making process.

9.11 SELF ASSESSMENT QUESTIONS

1. Explain the various economies of scale?
2. Explain the determinants of cost function?
3. Explain the econometric method of estimating cost function? Why is this method more popular than the other two methods (accounting and engineering) estimation costs?
4. What are the common problems you encounter while attempting to derive empirical cost functions from economic data?

5. The total cost function for a manufacturing firm is estimated as

$$C=128+6Q+2Q^2$$

Determine the optimum level of output Q to be produced?

6. Suppose that for a XYZ corporation's total cost function is as follows

$$TC=300+3Q+0.02Q^2$$

Where TC is the total cost, Q is the output.

- a. What is the corresponding fixed cost function, average fixed cost function, and variable cost function, average variable cost function?
 - b. Calculate the average total cost function and marginal cost function.
7. Based on a consulting economist's report, the total and marginal cost functions for an ABC company are

$$TC = 200 + 5Q - 0.04Q^2 + 0.001Q^3$$

$$MC = 5 - 0.08Q + 0.003Q^2$$

The president of the company decides that knowing only these equations is inadequate for decision making. You have been directed to do the following.

- a. Determine the level of fixed cost (if any) and equations for average total cost, average variable cost, and average fixed cost.
 - b. Determine the rate of output that results in minimum average variable cost.
 - c. If fixed costs increase to ₹ 500, what output rate will result in minimum average variable cost?
8. Given the total cost function for Laxmi Enterprises Co.

$$TC = 100Q - 3Q^2 + 0.1Q^3$$

- a. Determine the average cost function and the rate of output that will minimize average cost.
9. Determine the marginal cost function and the rate of output that will minimize marginal cost.

9.12 FURTHER READINGS

Adhikary, M. (1987). *Managerial Economics* (3rd ed.). Khosla Publishers, Delhi.

Maddala, G. S., & Miller, E. M. (1989). *Micro Economics: Theory and Applications*. McGraw-Hill, New York.

Maurice, S. C., Smithson, C. W., & Thomas, C. R. (2001). *Managerial Economics: Applied Microeconomics for Decision Making*. McGraw-Hill Publishing.

Mote, V. L., Paul, S., & Gupta, G. S. (2016). *Managerial Economics: Concepts and Cases*. Tata McGraw-Hill, New Delhi.

Dholakia, R., & Oza, A. N. (1996). *Microeconomics for Management Students*. Oxford University Press, Delhi.

Lewis, W.C., Jain, S. K., & Peterson, H. C. (2005). *Managerial Economics* (4th ed.). Pearson.

ignou
THE PEOPLE'S
UNIVERSITY

Block 4**PRICING DECISIONS**

Unit 10: Market Structure and Barriers to Entry	227
Unit 11: Pricing Under Perfect Competition and Pure Monopoly	251
Unit 12: Pricing Under Monopolistic and Oligopolistic Competition	269
Unit 13: Pricing Strategies	287

BLOCK 4 PRICING DECISIONS

Block 4 introduces the different market structures and their implication for pricing behavior of firms. In **unit 10**, market structure and the significant concept of entry barriers is discussed. It examines the reasons why certain entry barriers are 'natural' i.e. determined by the nature of the industry and why certain entry barriers are created by managers themselves in order to maximize current or even long-term profitability. **Unit 11** dwells on the extreme types of markets structures i.e. perfect competition and monopoly and rationalizes them as benchmarks. The outcomes achieved under these markets are used to compare outcomes achieved in other markets that lie in between these two market structures. Although both these structures are difficult to find in practice,, they are useful from the public policy point of view. **Unit 12** examines monopolistic competitions and oligopoly and briefly introduces the tools of game theory. This is done in an extremely elementary manner, so as not to detract from the main focus of the unit. The concept of market concentration and its measurement is also discussed. Finally **Unit 13** explores the various methods of pricing followed in different kinds of markets introduced in the previous units. It explains the concepts of price discrimination and its nuances and argues that to engage in price discrimination, the firm must have some control over price. Other interesting pricing techniques viz. bundling, peak load pricing, two part and multipart pricing along with a number of examples are developed.

ignou
THE PEOPLE'S
UNIVERSITY

UNIT 10 MARKET STRUCTURE AND BARRIERS TO ENTRY

Objectives

After going through this unit, you should be able to:

- **understand** the concept of market structure and the impact it has on the competitive behaviour of the firms;
- **classify** different types of market structures;
- **analyze** the factors that influence the pricing decisions of a firm;
- **identify** the barriers to entry of firms in the market.

Structure

- 10.1 Introduction
- 10.2 Classification of Market Structures
- 10.3 Factors Determining the Nature of Competition
- 10.4 Barriers to Entry
- 10.5 Strategic Entry Barriers—A Further Discussion
- 10.6 Pricing Analysis of Markets
- 10.7 Summary
- 10.8 Key Words
- 10.9 Self-Assessment Questions
- 10.10 Further Readings

10.1 INTRODUCTION

One of the most important decisions made by managers is setting the price of the firm's product. If the price set is too high, the firm will be unable to compete with other suppliers in the market. On the other hand, if the price is too low, the firm may not be able to earn a normal rate of profit. Pricing is thus a crucial decision area, which needs much of managerial attention.

In this unit we will examine the factors that govern this key decision area. Traditional economic theory explains this in term of demand and supply functions. According to traditional analysis, firms aim towards maximization of profits. The interplay of demand and supply in the market determines the price, which is often referred to as equilibrium price.

There are, however, many other factors that influence the pricing decision of a firm. These are – the number of firms in the industry, the nature of product, and the possibility of new firms entering the market and so on. In this unit you will understand more about some of the crucial factors that operate in the market place. In the process, you should gain valuable insights into the

operations of firms under different market structures, which are more typical of the existing real world situations.

10.2 CLASSIFICATION OF MARKET STRUCTURES

The structure of a market depicts the existence of firms in a particular market and to what extent the firms constituting a specified market are functionally interrelated to each other. The term 'market structure' refers to the degree of competition prevailing in that particular market. The power of an individual firm to control the market price by changing its own output determines the degree of competition and this power varies inversely with the degree of competition. The higher the degree of competition, the less market power the firm has and vice-versa. Market power is generally thought to be the ability of the firm to influence price.

A firm behaves according to its policies and practices regarding price, output decisions etc. The firm's performance is an indicator of its outcome or results of its conduct. The whole concept explains the Structure-Conduct-Performance (S-C-P) hypothesis. Hence in microeconomics theory, this hypothesis states that the performance of a firm is determined by its conduct, which in turn is determined by the structure of the market in which it is operating. The performance and the conduct of a firm vary from market to market. If the market is highly competitive then the performance and conduct of the firm is different as compared to that of the market with little or no competition. For example, pricing behaviour of firms in the fast moving consumer goods (FMCG) sector where there are a large number of rivals is very different from the pricing in the Airline industry where there are fewer firms.

Pricing decisions are affected by the economic environment in which the firm operates. Managers must, therefore, make their decisions to the specific market environment in which their firms operate. The central phenomenon in the functioning of any market is competition. Competitive behaviour is moulded by the market structure of the product under consideration. Since the decision-making environment depends on the structure of the market, it is necessary to have a thorough understanding of this concept.

The structure of a particular market plays an important role in defining the determinants that affect these market structures. Determinants like price, product differentiation etc. are affected by the competitive structure of the market. The classification of markets in terms of their basic characteristics helps identify a limited number of market structures that can be used to analyze decision-making. The four characteristics used to classify market structures are: i) Number and size distribution of sellers, ii) Number and size distribution of buyers, iii) Product differentiation and iv) Conditions of entry and exit.

i) Number and size distribution of sellers

The firm's ability to affect the price and the quantity of a product supplied to the market is related to the number of firms offering the same product. If there are a large number of sellers, the influence of any one firm is likely to be less. Consider the number of firms selling fruits and vegetables in your locality. It is unlikely that any one of them will exercise a great influence over price. On the contrary, if there are only few sellers in the market, an individual firm can exercise greater control over price and total supply of the product. Considering this fact the number of firms can be classified into large, few, two and one.

ii) Number and size distribution of buyers

Markets can also be characterized by the number and size distribution of buyers, where there are many small buyers of a product and all are likely to pay about the same price. Consider a big firm in a city. For example, A Steel manufacturing Company in Jamshedpur is a large and perhaps the only firm in the area. The Steel manufacturing Company will thus be able to exercise considerable influence on the price at which it buys inputs from suppliers in the area. Similarly, An Automobile Company in Gurgaon is one of the large automobile manufacturers and has considerable influence over the price at which it buys inputs such as glass, radiator caps and accessories from other suppliers located in the region. Both the Automobile and Steel manufacturing Companies are firms that are said to have 'monopsony' power in their buying decisions. However, if there are a large number of buyers they will be unable to demand lower prices from sellers. One reason why large firms are able to negotiate lower prices is because of large volume purchases.

iii) Product Differentiation

If the products competing in the market are not identical or homogeneous, they are said to be differentiated and hence 'product differentiation' exists in the market. Product differentiation is a fact of life and there is some amount of differentiation for almost all products that we buy in markets. For example, ingredients in different soaps could be different as can be the packaging, advertising etc. Even seemingly homogeneous goods such as apples and bananas are at present differentiated on the basis of the orchards where they have been grown and the way these are marketed. Wheat is a good example of a product that can be considered undifferentiated. The degree of substitutability or product differentiation is measured by cross-elasticity of demand between two competing products. This feature was explained in unit 5. Products can be classified into perfect substitutes or homogeneous products, close substitutes like soaps of different brands, remote substitutes like radio and television and no substitutes like cereals and soaps. Further, perfect substitutes for one consumer may not be so for another. For example, Rahul may feel that products C and P are perfect

substitutes while Sachin may have a strong brand preference for product P. Product differentiation is a basis for a lot of advertising that is seen in the media where the focus is to create a strong brand preference for the product being advertised.

iv) Conditions of Entry and Exit

Entry or exit of firms to an industry refers to the difficulty or ease with which a new firm can enter or exit a market. In short run, where the capital of firms is fixed, entry and exit does not make much difference. Ease of entry and exit is however a crucial determinant of the nature of a market in the long run. When it is difficult for firms to enter the market, existing firms will have much greater freedom in pricing and output decisions than if they had to worry about new entrants. Consider a pharmaceutical firm that has a patent on a particular drug. A patent is an exclusive right to market the product for a given period of time, say 12 years. If there are no close substitutes to that drug, the firm will be free from competition for the duration of the patent. Thus the barriers to entry in the market for this drug are high. Similarly, since Indian Railways, is a public monopoly no new entrant can enter the market. An American multinational technology corporation which produces computer software and related services too has been able to create substantial entry barriers in the market making it difficult for new firms to enter in the market. On the other hand, retail outlets and the restaurant business witness several new firms entering the market periodically, implying that entry barriers are relatively low.

Based on the above characteristics markets are traditionally classified into four basic types. These are Perfect Competition, Monopoly, Oligopoly and Monopolistic Competition.

Perfect competition is characterized by a large number of buyers and sellers of an essentially identical product. Each member of the market, whether buyer or seller, is so small in relation to the total industry volume that he is unable to influence the price of the product. Individual buyers and sellers are essentially price takers. At the ruling price a firm can sell any quantity. Since there is free entry and exit, no firm can earn excessive profits in the long run.

Monopoly is a market situation in which there is just one producer of a product. The firm has substantial control over the price. Further, if product is differentiated and if there are no threats of new firms entering the same business, a monopoly firm can manage to earn excessive profits over a long period.

Perfect Competition and Monopoly are discussed in more detail in unit 11.

Monopolistic competition a term coined by E. M. Chamberlin implies a market structure with a large number of firms selling differentiated products. The differentiation may be real or is perceived so by the customers. Two brands of soaps may just be identical but perceived by the customers as different on some fancy dimension like freshness. Firms in such a market

structure have some control over price. By and large they are unable to earn excessive profits in the long run.

Since the whole structure operates on perceived product differentiation, entry of new firms cannot be prevented. Hence, above normal profits can be earned only in the short run.

Oligopoly is a market structure in which a small number of firms account for the whole industry's output. The product may or may not be differentiated. For example, only 5 or 6 firms in India constitute 100% of the integrated steel industry's output. All of them make almost identical products. On the other hand, passenger car industry with only three firms is characterized by market differentiation in products. The nature of products is such that very often one finds entry of new firms difficult. Oligopoly is characterized by vigorous competition where firms manipulate both prices and volumes in an attempt to outsmart their rivals. No generalization can be made about profitability scenarios

We will discuss Monopolistic Competition and Oligopoly in detail in unit 12.

It must also be noted that these market structures can be classified in only two fundamental forms – **Perfect Competition** and **Imperfect Competition**. Under this classification, Monopoly, Oligopoly and Monopolistic Competition are treated as special cases of markets, which are less than perfect. Thus, these forms illustrate the degree of imperfection in a market by using the number of firms and product differentiation as basic criteria. Table 10.1 provides a ready reference for different types of markets based on their characteristics.

Table 10.1 Classification of markets based on their characteristics

Type of market structure	Basis of Distinction			
	Number of independent sellers	Seller concentration	Product differentiation	Condition of entry
Perfect or Pure competition	Large	Non-existent	Homogeneous product	Free or easy
Monopolistic competition	Large	Non-existent or low	Products are close substitutes	Free or easy
Oligopoly	Few	Medium or high	Products may be homogeneous or close substitutes	Difficult
Duopoly	Two	High	Products may be homogeneous or close substitutes	Very difficult or impossible
Monopoly	One	Very high	Remote substitutes	Barred or impossible

Activity 1

Suppose you are working in a company dealing with fast moving consumer goods. Classify the products of your company and its competitions under the type of competition it operates in and why?

.....

.....

.....

.....

.....

.....

10.3 FACTORS DETERMINING THE NATURE OF COMPETITION

We have already seen that the number of firms and product differentiation are extremely crucial in determining the nature of competition in a market. It has been tacitly assumed that there are a large number of buyers. What would happen if there are several firms producing standardized product but only one buyer?

Obviously, the buyer would control the price, he will dictate how much to buy from whom. The entire price-volume decision takes on a different qualitative dimension. Similarly, product features and characteristics, the natures of production systems, the possibility of new entrants in a market have profound impact on the competitive behaviour of firms in a market. The 'entry' of new firms has special relevance in business behaviour which we discuss in the next section and deal with other issues in the present one.

Effect of Buyers

We have already referred to the case where there is only one buyer. Such a situation has been referred to as **monopsony**. For example, there are just six firms in India manufacturing railway wagons all of which supply to just one buyer, the Railways. Such a situation can also exist in a local labour market where a single large firm is the only provider of jobs for the people in the vicinity. More frequently encountered in the Indian markets is a case of a few large buyers, defined as **oligopsony**. The explosive industry which makes detonators and commercial explosives has three major customers: Coal India Ltd. (CIL), Department of Irrigation and various governmental agencies working on road building activities. Of these, just one customer, CIL takes nearly 60% of the industry's output. There are about 10 firms in the industry, which negotiate prices and quantities with CIL to finalize their short-term plans.

Most industries manufacturing heavy equipment in India are typically dominated by a few manufacturers and few buyers with the Government being the major buyer. Price and volume determination in such products often takes the form of 'negotiation across the table' rather than the operation of any market forces. Since the members in the whole market inclusive of buyers and sellers are not many, very often they know each other. In other situations, like the consumer goods, firms have no direct contact with their customers.

Production Characteristics

Minimum efficient scale (MES) of production in relation to the overall industry output and market requirement sometimes plays a major role in shaping the market structure. MES is the minimum scale of output that is necessary for a firm to produce in order to take advantages of economies of scale. For example, the minimum efficient scale for an automobile firm is very high. This is intuitively appealing because if only 100 cars are produced in a capital intensive automobile plant, the average costs will be high, while a larger volume of cars will allow the fixed costs to be spread over a number of cars, thus reducing average costs and increasing the minimum efficient scale. MES for a service firm such as a travel agent will accordingly be relatively small.

The reason why there are no more than say, around 10 integrated steel plants even in an advanced country like the U. S. A. can be partly explained by economies of scale and thus MES. Since the minimum economic size of such a steel plant is a few million tonnes, the entire world steel industry can have no more than 100 efficient and profitable firms. Thus every country has only a handful of steel plants. On the other hand, when one comes to rolling mills which take the steel billets or bars as input, the minimum efficient size comes down considerably, and given the existing demand, several firms can be seen to operate.

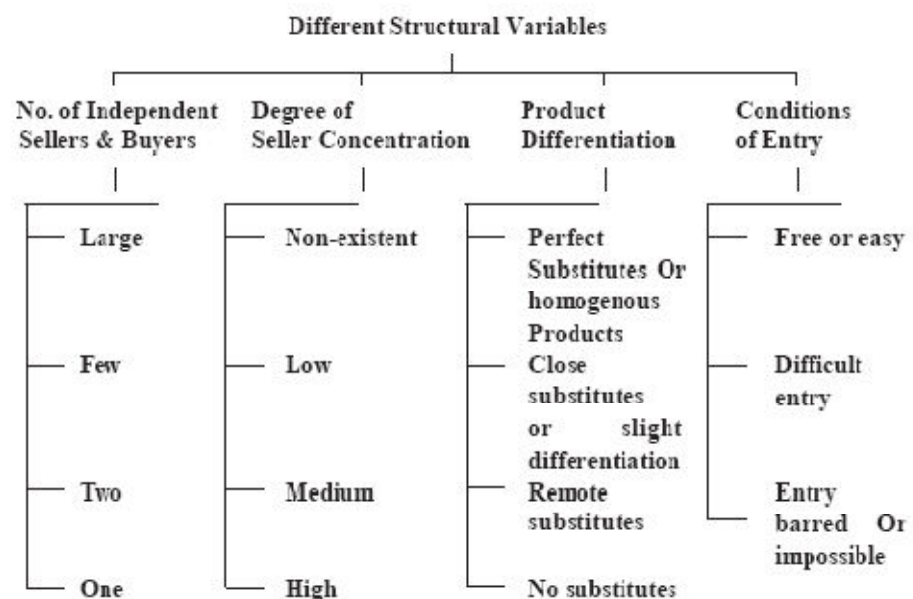
Further, the minimum size does not remain constant but changes drastically with technological advancements. When technical changes push up the economic size of a plant, one notices that the number of firms decline over time. This can be noticed in some process industries like synthetic fibre. Conversely, technological innovations may make it possible for smaller sized plants to economically viable. In such a case a lot of new entrants come and soon the market becomes highly competitive as has happened in the personal computer industry in India.

Apart from minimum plant size, factors like the availability of the required raw material, skilled labour etc. can also mould market structures. Likewise, enough skilled people are not available to work on the sophisticated machines. These factors sometimes restrict output and push up prices even though adequate market potential for expansion exists.

We have already stated that product differentiation is an important market characteristic because it indicates a firm's ability to affect price. If a firm's product is perceived as having unique features, it can command a premium price and the firm is said to possess market power. For example, the Luxury Cars have come to be regarded as the ultimate in automobile luxury and therefore commands a high price. Consumers are willing to pay that premium for the product. The degree of competition faced by Luxury Cars is thus very low. One could also consider the example of railway services in India. Indian Railway Catering and Tourism Corporation (IRCTC) has monopoly in railways so people are just dependent on IRCTC for railway ticket bookings. Thus, the market for railway services is not competitive in the sense that only one corporation provides the service. Although this monopoly is due to government restrictions in India which is going to change in the coming years. On the other hand, for a product like soap or detergents, there are many firms producing a large variety of substitutable products. Therefore, one notices more competition, in the detergent market. The physical characteristics of a product can also influence the competitive structure of its market. If the distribution cost is a major element in the cost of a product, competition would tend to get localized. Similarly, for perishable products, the competition is invariably local.

Conflict between physical characteristics and minimum economic size

An interesting question arises in the case of a product like cement. For reasons of minimizing the transport costs on raw materials, most cement plants in the country are located near mine sites. A large efficient plant near a mine site can manufacture cement at the optimum cost, but the local demand is never large enough. If such a plant has to sell in far away markets (from Gujarat to Kerala, for example) the transport costs can be quite high. Customers located in such areas will always buy cement at a much higher price. The government partly offsets this by using the mechanism of levy price which is the same throughout the country.



Market selection: Entry and Exit

Market selection process includes firm's entry, then its survival and finally the exit process. The selection and expansion depends how efficient the firm is. The efficient firms enter and the inefficient ones exit.

Conditions of Entry: The entry of a new firm in an industry or a market depends on the ease with which it can enter. If we see the long-term perspective, the number of firms and the degree of seller concentration depends on the conditions of entry. In case of free entry, the number of sellers is large in number and in case of restricted entry, the number of sellers tend to reduce. In the long run the degree of competition depends on the condition of entry. A new entrant could bring with it the following advantages.

- Provides new goods and services,
- Changes the balance between different sectors,
- Comes with new technological and managerial techniques,
- Increases opportunities.

Factors determining conditions of entry

The following are some of the factors that determine the structure of any market. This list is not meant to be exhaustive, but is likely to cover a large part of real world situations.

- Legal barriers
- Initial capital cost
- Vertical integration
- Optimum scale of production
- Product differentiation

Legal barriers: Almost all countries have their set of rules and regulations. Patent law is one such regulation, which promotes and protects the interests of inventors and innovators. Under this law, no firm other than the patent holder or the licensed firm is allowed to make use of the process. India has its own legal barriers and it had certain laws like Industrial Licensing Regulation and Reservation of products, which restrict entry and thus protect the incumbent firm from competition.

Initial capital cost: For industries producing basic inputs like coal, steel, power etc., the initial capital cost is quite high. Therefore, it becomes difficult for new entrepreneurs to enter. In industries where the capital requirement is high, the market is dominated by a few firms, whereas for industries such as non-durable consumer goods, the initial capital cost is less and therefore the number of firms in the market can be quite large.

Vertical integration: A vertically integrated firm is one that produces raw material i.e. an intermediate product as well as the final product. Examples of vertically integrated firms in India are integrated steel plants such as SAIL and TISCO and DMart a supermarket chain. Entry in this case is restricted to limited producers as here the existing producer produces raw material or an intermediate product along with the final product. New entrants will find that their capital requirements are high and hence it will not be easy for them to enter the market.

Optimization: Optimum scale of production means the scale of output at which the long run average cost of production is minimum. As defined earlier this is the minimum efficient scale of production for the firm. If the optimum scale of output for any product is quite large and if the total market can be efficiently served by a few firms, the new entrants will find it difficult to enter such markets. Examples of such markets are electricity generation and aircraft production.

Product differentiation: New entrants face difficulty to enter the market where the products are highly differentiated. Consider an example of a popular ready to eat breakfast cereal company. It is the market leader and produces more than 40 different kinds of cereal ranging from the ordinary corn flakes to granola flakes and muesli. With such a wide variety, new entrants find it difficult to compete with it for shelf space in retail outlets which is crowded with its products. By implementing such widespread product differentiation, it has managed to increase the cost of entry for potential entrants in the market.

Related to entry conditions is the concept of entry barriers. Any manager is concerned about his firm's market share and thus threat to its competitive position. By establishing an entry barrier a firm not only preserves its market share but could also increase it. This is perhaps the most interesting aspect of market structure and its analysis. Such attempts are made everyday by managers and are widely visible in the environment around us. An example of an entry barrier is advertising expenditure by firms. Think about the enormous advertising spend of soft drink firms such as C and P and examine whether it is possible for a new entrant to try and compete with such large existing brands even if it come up with an equally good soft drink. We will study this feature of markets in detail now.

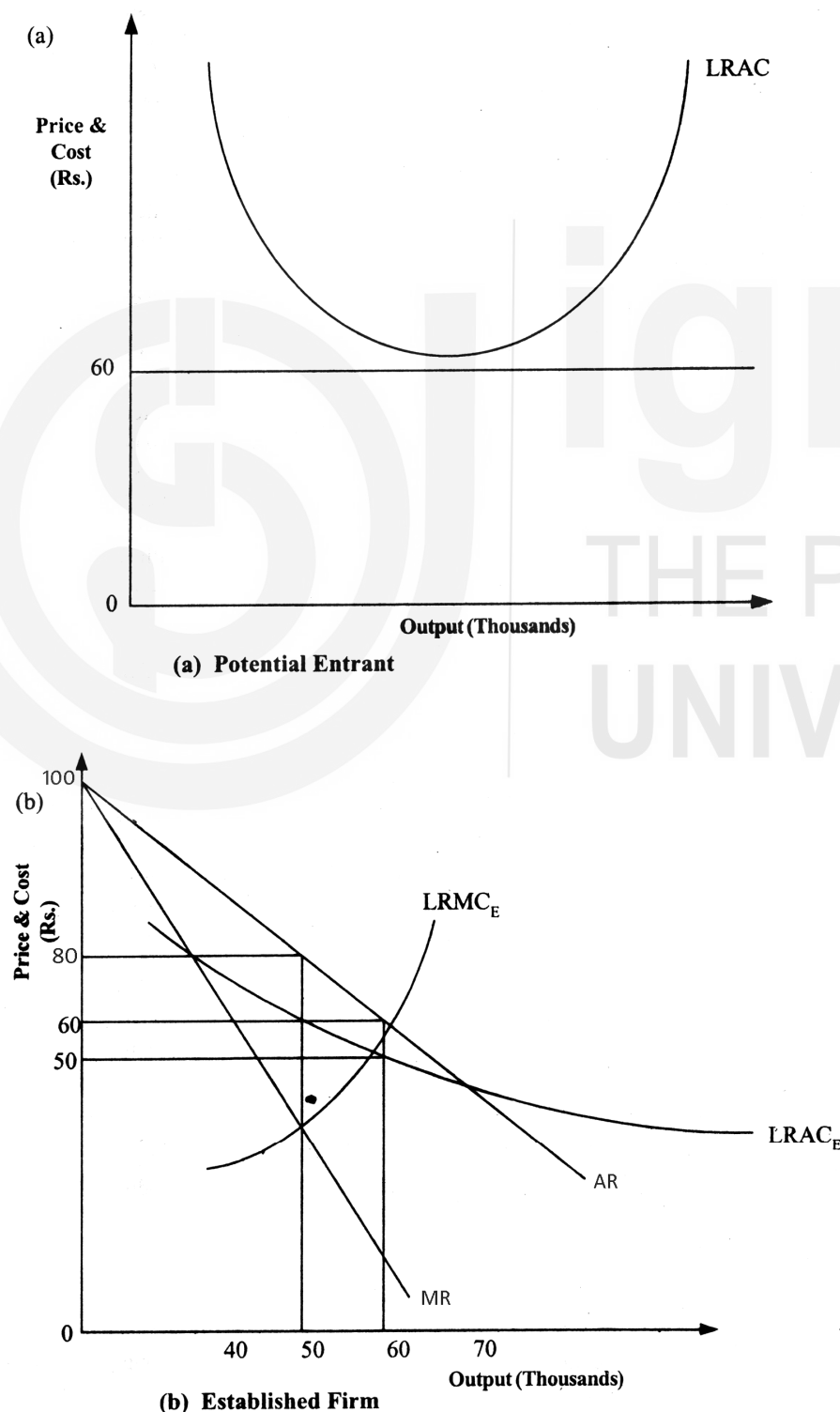
A barrier to entry exists when new firms cannot enter a market. There are many types of barriers, which become sources of market power for firms. *Entry barriers* can be broadly classified as: *Natural barriers, Legal Barriers and Strategic Barriers*.

Natural barriers: Economies of scale create a natural barrier to the entry of new firms and it occurs when the long run average cost curve of a firm decreases over a large range of output, in relation to the demand for the product. Due to the existence of substantial economies of scale, the average cost at smaller rates is so high that the entry is not profitable for small-scale firms. This results in existence of natural monopoly. Power generation, Aircraft manufacturers, Railways, etc. are examples of natural monopolies. You should keep in mind that technological progress often undermines the

natural monopoly character of certain industries. This has happened in telecommunications, which not very long ago used to be considered a natural monopoly.

Legal barriers: Patents, as discussed above, are an example of a legal entry barrier. Industrial licensing that used to be common in India in the 1970s and 80s is another example of such a barrier. By giving a license to a firm the government provided exclusive rights to that firm or a few firms to produce. This restricted the number of players in the market through industrial licensing, thus creating a legal entry barrier.

Figure 10.1: Entry Limit Pricing



Strategic barriers: Such barriers exist exclusively due to the strategic behaviour of existing firms. Managers undertake investments to deter entry by raising the rival's entry costs. To bar or restrict the entry of a new entrant, an established firm may change price lower than the short-run profit-maximizing price. This strategy is known as entry limit pricing. The entry limit pricing depends on established firm taking a cost advantage over potential entrants. The established firm must have a long run average cost curve below that of the other firm in order to lower its price and continue to make an economic profit.

For example, established firm lowers its price below profit-maximizing level. Figure 10.1 shows demand and marginal revenue curves for an established firm and also the firm's long run average (LRAC) cost and marginal cost (MC) curves as $LRAC_E$ and $LRMC_E$.

To maximize profit, the firm produces 50,000 units of output when $MR=MC$ and fix a price of Rs. 100 from the demand curve. Therefore the firm's profit becomes:

$$P = (\text{Rs. } 100 - \text{Rs. } 80) * \text{Rs. } 50,000 = \text{Rs. } 10,00,000$$

Suppose the established firm sets the prices at Rs. 90 for say 70,000 units of output, the new entrant would not be able to cover the average cost as it would be making loss. The economic profit of the established firm now would be:

$$EP = (\text{Rs. } 90 - \text{Rs. } 80) * \text{Rs. } 70,000 = \text{Rs. } 7,00,000$$

Though this profit is less than the original profit but if we look at the practical point, it is found that even if the established firm incurs a loss, the sales of the firm can be increased in the future regarding the difficulties posed for the new entrant. The lower profit would be higher had the new firm entered the market and would have taken away some share of the sales from the established firm. This example shows that **entry-limit pricing** is not feasible without the cost advantage.

Building Excess Capacity: Another way to restrict the entry is to build and maintain excess capacity over and above the required amount. This poses a threat to the new entrant deliberating the fact that the established firm is prepared to increase the output and lower the price if and when entry occurs. The excess capacity can be built up easily as it takes a longer time for the new entrant to build a factory of such capacity. This type of barrier is also known as **capacity barrier to entry**.

Producing Multiple Products: Economies of scope arise when cost of producing two or more goods together is less costly than producing the two goods separately. The process goes on and becomes cost effective as more goods are produced. This acts as entry deterrent for new firms.

New Product Development: Producing substitutes for its own product in the market can discourage the entry for the new firms. For example a consumer good company producing different types of soaps targeted to different customer base. The more the number of substitutes, the lower and more elastic is the demand for any given product in the market. This makes the entry of new firm more difficult.

Take the case of Microsoft. Why does every other personal computer (PC)/laptop that one comes across is Microsoft windows compatible. The PC cannot work without windows. By developing industry level standards, Microsoft has created ‘high switching costs’ in an attempt to create entry barriers.

Activity 2

Given below is the list of some industries. Indicate in column 3 whether the entry barriers are high or low. Give reasons in column 4.

S.No.	Name of the Industry	Entry Barriers	Reasons
1.	Software		
2.	Hardware		
3.	Oil-field chemicals		
4.	CNC machine tools		
5.	Breakfast cereals		
6.	Aluminium		
7.	Ball-point pens		
8.	Television Sets		
9.	Cement		
10.	Chocolates		

10.5 STRATEGIC ENTRY BARRIERS — A FURTHER DISCUSSION

No one likes competition and companies with a leading position in a market will go to considerable lengths to keep out likely new opponents. Although all companies strive to develop one form of competitive advantage or another, relatively few are persistently successful over long periods. Innovative activity is almost always followed by waves of imitation and relatively few first movers are able to maintain their initial market position.

One of the Indian instant messaging apps founded in 2016 was so popular for its funny and cool stickers that every month million users got on it. It was one of the youngest startup which became unicorn in short span of time with a valuation of over \$ 1 billion. But the business could not sustain competition from foreign instant messaging apps. It had to shut down its operations

in early 2021. despite being one of the major innovations of Indian communication industry. The simple truth is that most large-scale expenditures designed to create competitive advantage are unlikely to realize a return unless that advantage can be sustained.

Economists think about this problem as one of creating, or strategically exploiting, barriers to entry or mobility barriers. Entry barriers, as defined above are structural features of a market that enable incumbent companies to raise prices persistently above costs without attracting new entrants (and, therefore, losing market share). Entry barriers protect companies inside a market from imitators in other industries. Entry barriers give rise to persistent differences in profits between industries.

Although different commentators produce different lists, almost all sources of entry barriers fall into one of the three following categories: product differentiation advantages, absolute cost advantages, and scale-related advantages. Product differentiation arises when buyers distinguish the product of one company from that of another and are willing to pay a price premium to get the variant of their choice. Such differences become entry barriers whenever imitators, whether they be new entrants or companies operating in other niches of the same market, cannot realize the same prices for an otherwise identical product as the incumbent. On the face of it, it is hard to understand how this might come about since consumers will (surely) always prefer the lower-priced variant of two otherwise identical products.

However, if it is costly for consumers to change from purchasing one product to purchasing another, then prices for otherwise identical products can differ for long periods of time.

Economists call costs of this type switching costs and business managers always try to create switching costs by locking consumers into their product. Habit formation is an obvious source of switching costs and many marketing campaigns are designed to reinforce the purchasing patterns of existing customers and raise their resistance to change. Further, many consumers sink costs into gathering information about new products and, once they have made a choice that satisfied them, they are likely to resist making further investments

Both sources of switching costs are often reinforced by the use of brand names to help consumers quickly find familiar products. The value of these labels depends, of course, on the size of the switching costs that they help to sustain. Finally, switching costs also arise when consumption involves the purchase of highly specific complementary products that lock consumers into existing purchasing patterns. Buyers of a foreign company's mainframes often found that the large costs of rewriting software and recording data

dwarfed price or performance differences that might otherwise have induced them to switch to one of its foreign rivals.

Absolute cost advantages arise whenever the costs of incumbent companies are below those of new rivals and they enable incumbents to under-cut the prices of rivals (by an amount equal to the cost disadvantage) without sacrificing profits. There are many sources of absolute cost advantages. Investments in R&D and learning-by-doing in production can be important in many sectors and they can occasionally be protected by patents. Similarly, privileged access to scarce resources (such as deposits of high-quality crude oil, much sought after airport landing slots or the odd scientific genius) can open up substantial differences in costs between companies producing identical products. Many companies vertically integrate upstream to assure control over limited natural resources or downstream to assure access to the most valuable distribution channels, actions that can make entry anywhere in the value chain difficult.

Scale-related advantages create the most subtle form of entry barriers. They arise whenever a company's costs per unit fall as the volume of production and sales increases. Economies of scale in production (created by set up costs, an extensive division of labour, and advantages in bulk buying and so on) are the most familiar source of scale advantages but economies can also arise in distribution. One way or the other, the important implication of scale advantages is that they impede small-scale entry. If costs halve as production doubles, then a small entrant will have costs per unit twice as high as an incumbent twice its size. Since it is unlikely that such an entrant will be able to differentiate its product enough to justify a price difference of this size, it must either enter at a scale similar to that of the incumbent or not enter at all. Needless to say, this compounds its problems, since raising the finance to support a large-scale (and therefore much riskier) assault on a privileged market can be much more difficult than raising funds for a much more modest endeavor.

As stated above, few markets naturally develop entry barriers and, even when they do, very few incumbent companies rely on structural features of market alone to protect them. Whether it is creating or exploiting entry barriers, companies with profitable market positions to protect usually need to act strategically to deter entry. Although there are as many different examples of strategic entry deterrence, there are at least three types of generic strategies that companies typically employ: sunk costs, squeezing entrants and raising rival's costs.

Sunk costs: Displacing incumbents is possibly the most attractive strategy for an entrant to follow since, if successful, it enables the entrant both to enter a market and monopolies it. Somewhat more modestly, if an entrant can at least partially displace an incumbent, it will make more profit after entry than if it

has to share the market on a less equal basis. To deter entrants from following this strategy, an incumbent needs to lock itself into the market in a way that raises the cost to the entrant of displacing it. This usually requires the incumbent to make investments whose capital value is hard to recover in the event of exit. Sunk costs raise the costs of exit (and so make it that much harder for the entrant to force the incumbent out). Some incumbents do this by investing in highly dedicated, large-scale plant and equipment since this also enables them to reap economies of scale in production. These activities also have the additional benefit of creating product differentiation or absolute cost advantages.

Squeezing entrants: It is usually all but impossible to deter very small-scale entry and frequently it is not worth the cost. However, capable entrants interested in establishing a major position in a market are a much more serious threat and many entry-deterring strategies work by forcing entrants to enter at large scale while at the same time making this too expensive. Squeeze strategies usually build on scale economies that prevent small-scale entry by forcing entrants to incur even more fixed costs (say through escalating the costs of launching a new product by extensively advertising), which increases their minimum scale of entry. Further, if these fixed costs are also sunk then these activities also increase the risks associated with entry. The squeeze comes through actions that limit their access to customers, making the larger scale of entry much more difficult and expensive to realize than a more modest market penetration strategy might have been. This is often done by filling the market with more and more variants of the generic product, developing fighting brands closely targeted on the entrant's product or limiting access to retail outlets.

A simple glance at the shelves of most super markets will reveal many instances where the multiple brands of a single company (or a small group of leading companies) completely fill all the available space, leaving little or no room for an entrant (examples might include laundry detergents of India's largest fast-moving consumer goods company and ready to eat breakfast cereals of an American multinational food manufacturing company).

Raising rival's costs: Even when an incumbent is sure that it cannot be displaced by an entrant and it has managed to squeeze the entrant into a tiny niche of an existing market, entry can sometimes be profitable when the market is growing.

Indeed, market growth is an important stimulus to entry since it automatically creates room for the entrant without reducing the incumbent's revenues. However, most entrants have only modest financial support and any strategy that raises costs in the short run and slows the growth of their revenues may make it difficult for them to survive long enough to penetrate the market and turn a profit. One rather obvious strategy of this type is to escalate advertising and, indeed, this is a very frequent response to entry by incumbents.

Advertising is a fixed cost (which, therefore, disadvantages small-scale entrants) and it is often the case that what matters is the relative amount of advertising a company does rather than the absolute amount. An advertising war initiated by an incumbent that raised total market advertising but keeps the advertising shares of companies relatively constant will, therefore, raise the entrant's costs without raising its revenues. The interesting feature of this strategy is that an advertising war will also raise the incumbent's costs. What is more, investments in advertising are often sunk, meaning that they are likely to raise the exit costs of the incumbent is able to turn what, on the face of it, appears to be a disadvantage to its advantage because entrants are more adversely affected by an advertising war than the incumbent is. That is, some investments that incumbents make seem irrational because they raise costs without generating much, if any, additional revenues. When successful, however, they are justified by the fact that they protect existing revenue streams from entrants. This points to one of the most characteristic features of investments in entry deterrence: they do not generate net revenue so much as they prevent it from being displaced.

A company that successfully deters entry will have lower profits than a company that did not face an entry threat but that is not an interesting observation. What matters is that a company that successfully deters entry will preserve its profits while a company that has not been able to deter entry will see its market position, and the profits that it generates, gradually disappear.

10.6 PRICING ANALYSIS OF MARKETS

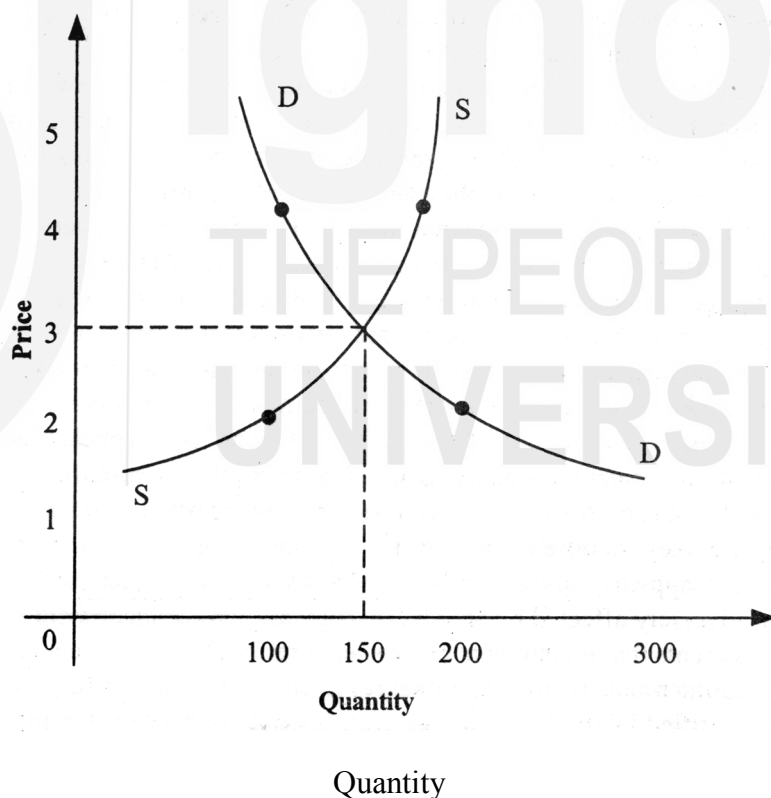
Pricing is an important function of all firms. Every firm is engaged in the production of some goods and/or services, incurring some expenditure to sell them in the market. It must, therefore, set a price for its product. It is only in extreme cases that the firm has no say in pricing its product because there prevails perfect competition in the market or the good has so much public significance that its price is decided by the government. Otherwise, in large number of cases, the individual producer plays the role in pricing her/his product.

The quantity supplied depends upon a number of factors. The law of supply takes into account its main determinant i.e price of the commodity. *The law of supply* states that supply of a commodity refers to the various quantities of the commodity which a seller is willing and able to sell at different prices in a market, at a point of time, ceteris paribus, other things remaining the same. Supply is related to scarce goods and not freely available goods. In this regard the role of demand and supply in determination of price is very important.

Table 10.2 Demand-Supply Schedule

Price	Demand	Supply
5	100	200
4	120	180
3	150	150
2	200	110
1	300	50

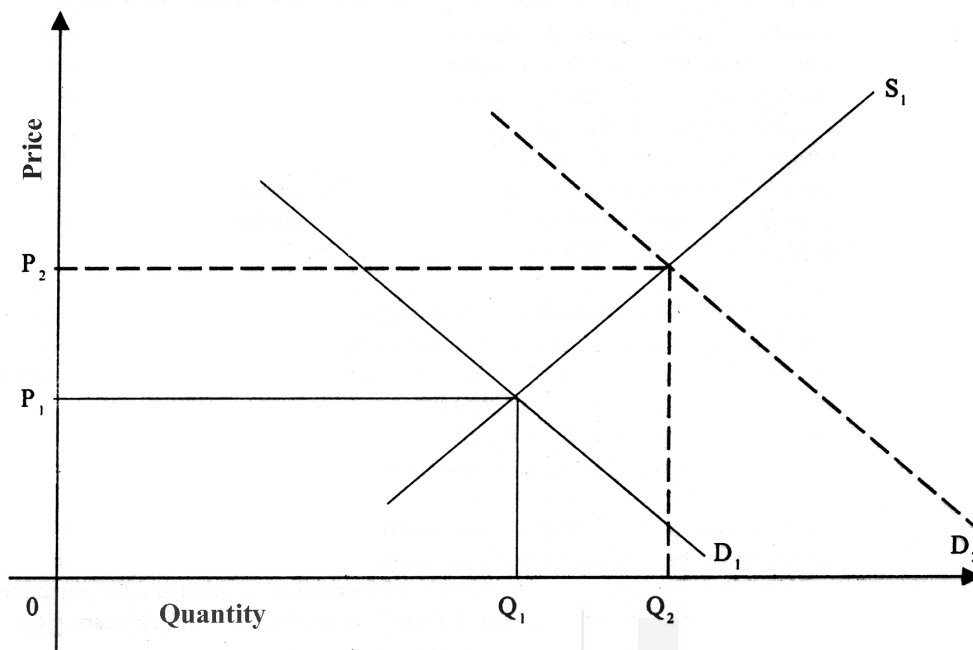
Setting the right price for its product is crucial for any firm in the market. This is because the price is such a parameter that it exerts a direct influence on the demand for and supply of the product and thereby on its sales and profit – the important yardsticks for the success or failure of the firm. If the price is set too high, the seller may not find enough customers to buy his/her product.

Figure 10.2: Demand-Supply curve

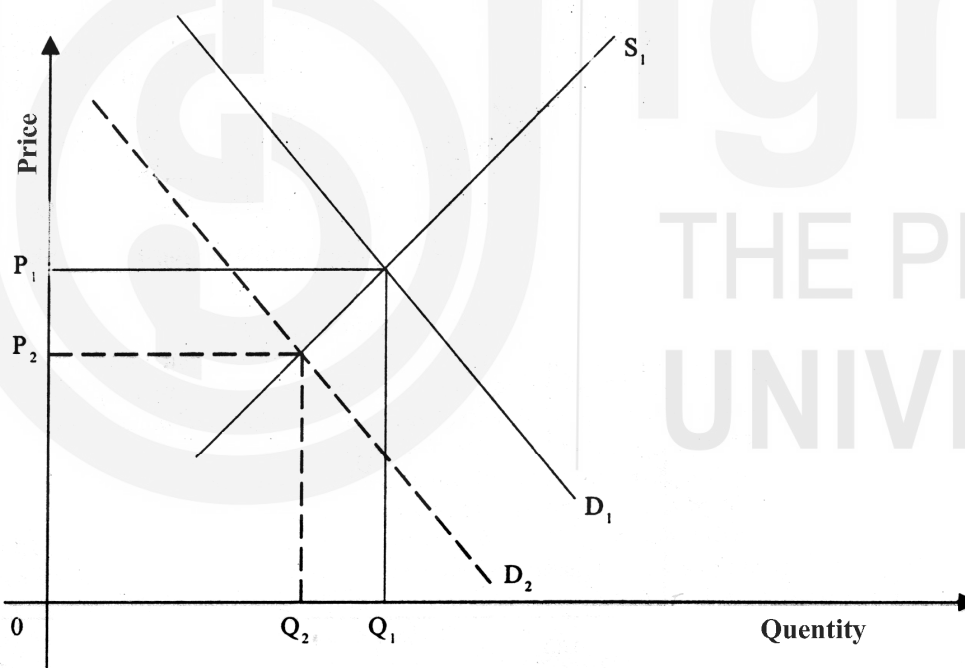
On the other hand, if the price is set too low, the seller may not be able to recover her/his costs. Further, demand and supply conditions vary over time and the managers must therefore review and reformulate their pricing decisions from time to time.

It is clear that the price of a product is determined by the demand for and supply of that product. Table 10.2 illustrates the demand and supply schedules of a product.

Figure 10.3: Effect of a change in demand on price and quantity



(a) Increase in Demand



(b) Decrease in Demand

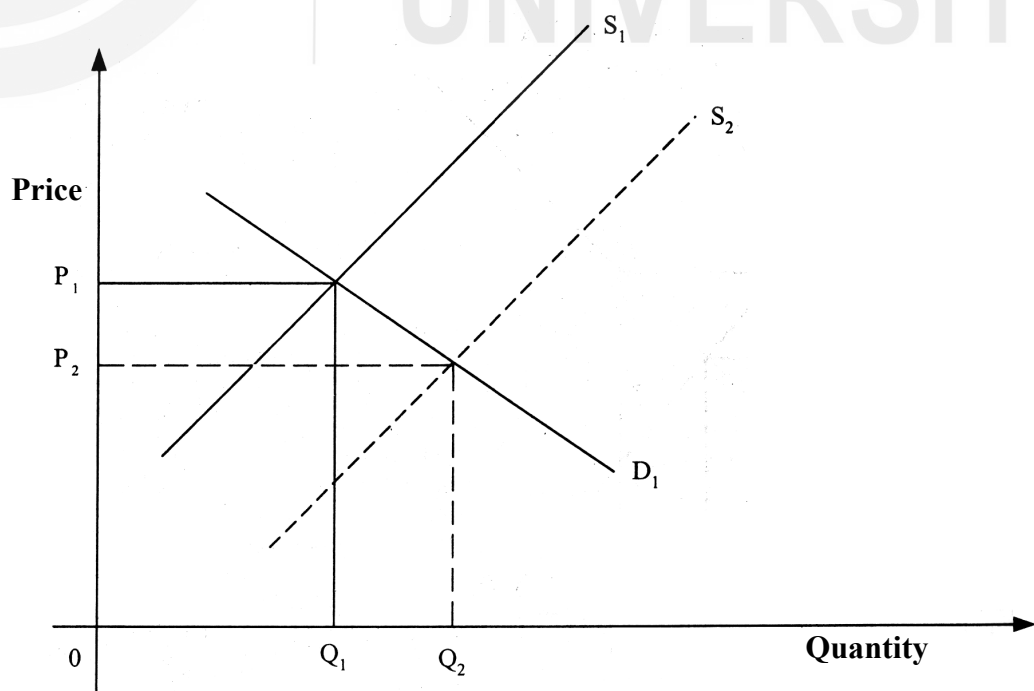
Let us assume that in the above example the market price, $P = 3$ and no other price prevails in the market (Figure 10.2). Because if $P = 5$, supply exceeds demand and the producers may not be able to find enough customers for their product. This would result into competition among the producers forcing them to bring down the price to 3. On the other hand, if $P = 1$, the demand exceeds supply which would give rise to competition among the buyers of the product, pushing the price up to 3. Therefore, at $P = 3$, demand equals supply, which is called equilibrium price. The equilibrium price is thus determined by the interaction of demand and supply.

We have seen in Block 2 that the demand for a good depends on a number of factors as does supply of a good. Therefore, the factors which affect either demand or supply are also determinants of price. A change in demand and/or supply would bring in a change in price. For instance, if the supply of a good is fixed, as shown in figure 10.3, the level of demand appears to determine the equilibrium price. In this case, the price is determined by the 'other factors' influencing the level of demand curve. An increase in demand from D_1 to D_2 , leads to an increase in equilibrium price from P_1 to P_2 and an increase in quantity from Q_1 to Q_2 [see figure 10.3 (a)]. Quite the opposite holds true in the event of a decrease in demand which is shown in figure 10.3(b).

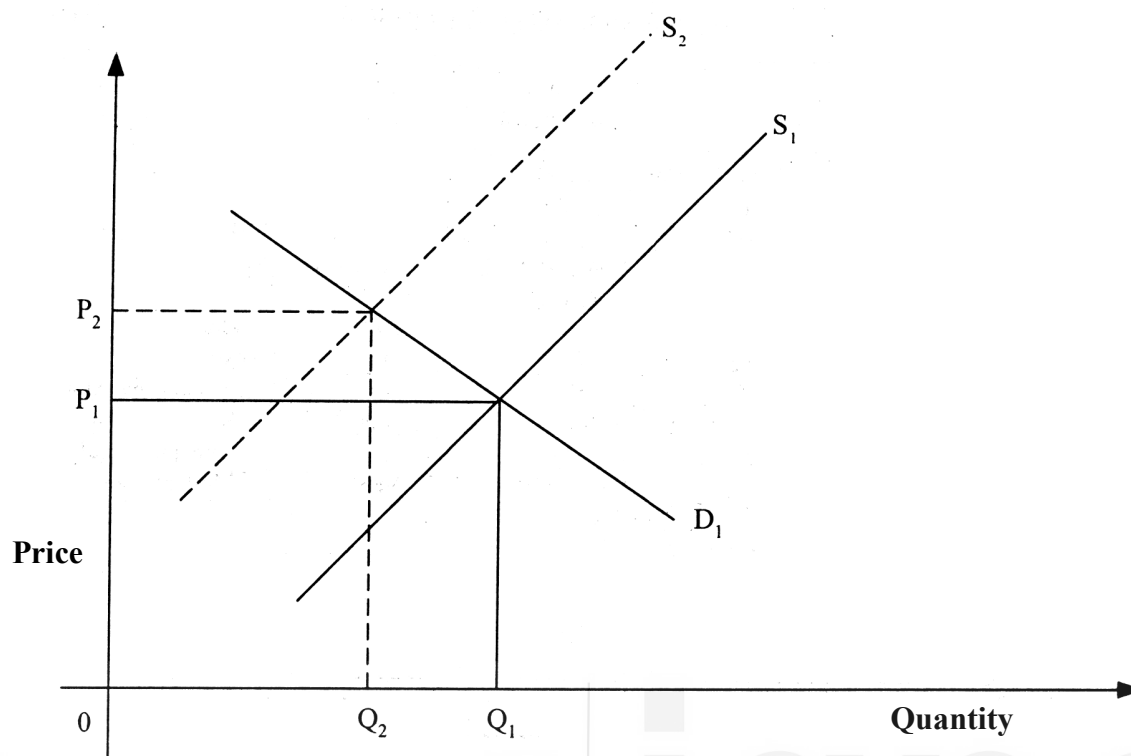
If the demand for a commodity is fixed, as shown in figure 10.4 the level of the supply curve determines the equilibrium price of the commodity. The equilibrium price would, therefore depend on the 'other factors' underlying the supply curve of the commodity. Figure 10.4 (a) shows that an increase in supply from S_1 to S_2 causes price to fall from P_1 to P_2 and the quantity to increase from Q_1 to Q_2 . Figure 10.4 (b) shows exactly the reverse case.

So far we have discussed the general equilibrium price which is determined by the interaction of demand and supply. However, the actual shapes of the demand and supply schedules depend on the structure of the product, market and the objectives of the firm. Thus market structure and firms' objectives also have a bearing on price.

Figure 10.4: Effects of a change in supply on price and quantity



(a) Increase in Supply



(b) Decrease in Supply

Since market structure influences price and different product groups fall under different market structures, pricing decisions depend upon market structure. For instance, automobile prices are set quite differently from prices of soap because the two products are produced by firms in different market structures.

Accordingly, in the subsequent units we shall discuss price determination under perfect competition and pure monopoly, and monopolistic competition and oligopoly (Units 12 and 13 respectively).

A large firm may produce a number of products, which are sold in variety of markets catering to the needs of different sections of the society. Let us take the example of FMCG, which produces products ranging from cosmetics to food products. Here comes the real task to be performed. At times it happens that price set for one of such products may affect the demand for the other product sold by the same firm. For example, the introduction of product R from an automobile company had an effect on the price of product 2 sold in the market.

Pricing of multiple products /a number of products produced by the same firm

It is difficult to set a price of multiple products but once it is set, the products make their own place in the market. Take the example of Fast Moving Consumer Goods Companies which are India's largest Packed Mass Consumption Goods (PMCG). The vision of the company is to meet the everyday needs of the people everywhere. Over the past years

these Company's have introduced somewhere around 100 brands, most of which have become household names in the country. The products vary from personal care products to beverages. A list of such products of FMCG Company is provided in Table 10.3. This example also gives an idea of product differentiation. In this case the price of each product is different because it caters to different segments of the market.

Table 10.3: List of Products by FMCG Companies

Beauty & Personal Care Products	Examples	Soaps and Detergents	Examples	Food, Beverages, Ice-cream, and Soups	Examples
ORAL CARE	<i>Tooth Paste Tooth Powder</i>	PERSONAL CARE	<i>Soaps Body Wash</i>	FOOD	<i>Flour, Salt,</i>
SKIN CARE	<i>Face Cream</i>				
	<i>Body Lotions</i>	FABRIC WASH	<i>Washing Powder & Detergents</i>	BEVERAGE	<i>Green Tea Tea granules Energy Drinks</i>
HAIR CARE	<i>Shampoos Hair Gel</i>	HOUSEHOLD CARE	<i>Utensil Cleaning Powder and liquid</i>	SOUPS & ICE- CREAM	<i>Ice cream cone Ice cream cup and soups</i>
BODY CARE	<i>Deodorant Perfumes</i>				
COLOUR COSMETICS	<i>Lipstick Nail Polish</i>				

Many Companies produce similar products under different brand names.

Activity 3

- 1) List five examples where the price of one product affects the demand for the other and vice-versa.

.....

.....

.....

.....

.....

- 2) List (any five) the name and product of the companies producing multiple products.

.....

.....

.....

.....

.....

10.7 SUMMARY

In this unit, we have made an attempt to understand the concept of market structure and the impact it has on the competitive behaviour of firms. Various competitive market situations were defined and broadly discussed. The number of firms and product differentiation are crucial determinants of the nature of competition in the market. The level of competition also gets influenced by number of sellers and buyers, buyers' behaviour, characteristics of product and production.

The pricing analysis of markets helps to understand how the equilibrium price is determined by the interaction of demand and supply. This forms the basis for analyzing the price-output decisions of firms under different competitive situations.

10.8 KEY WORDS

Market structure refers to the number and size distribution of buyers and sellers in the market for goods or service.

Perfect competition is a market structure where a large number of buyers and sellers deal in nearly identical products. Each is individually so small in relation to the total output that all members are 'price takers'.

Monopoly situation is characterized by just one producer of a product or service

Monopolistic competition is characterized by many sellers of a differentiated product.

Oligopoly situations have fewer sellers with or without the existence of product differentiation.

Product differentiation refers more to the differences in products as perceived by the customers than in real or technical difference in specifications.

Competition is the collective outcome of the forces generated within a given market structure (for a product or service) in combination with product characteristics, number of buyers, potential entrants and government policy.

Barriers to entry refer to the obstacles that impede the entry of new firms in an industry.

10.9 SELF-ASSESSMENT QUESTIONS

- 1) Classify the market structures based on certain factors and support your answer with the help of examples.
- 2) Discuss the different structural variables. Illustrate your answer with the help of examples.
- 3) Discuss the important technical barriers to entry.
- 4) Take the example of a hypothetical firm. Apply the strategic barriers to the firm and discuss.
- 5) The paperback books and the hardcover books are sold at different prices. Explain.
- 6) What are switching costs? Cite one example of a switching cost and examine how a firm can advantage from the existence of switching costs?

10.10 FURTHER READINGS

Mote, V. L., Paul, S., & Gupta, G. S. (2016). *Managerial Economics: Concepts and Cases*, Tata McGraw-Hill, New Delhi.

Maurice, S. C., Smithson, C. W., & Thomas, C. R. (2001). *Managerial Economics: Applied Microeconomics for Decision Making*. McGraw-Hill Publishing.

Dholakia, R., & Oza, A. N. (1996). *Microeconomics for Management Students*. Oxford University Press, Delhi.

UNIT 11 PRICING UNDER PERFECT COMPETITION AND PURE MONOPOLY

Objectives

- After going through this unit, you should be able to:
- **describe** the characteristics of pure/perfect competition and pure monopoly;
- **identify** the equilibrium conditions for a firm and the industry in a perfectly competitive situation;
- **examine** price-output decisions under pure monopoly;
- **analyze** the relevance of pure/perfect competition and pure monopoly.

Structure

- 11.1 Introduction
- 11.2 Characteristics of Perfect Competition
- 11.3 Profit Maximizing Output in the Short Run
- 11.4 Profit Maximizing Output in the Long Run
- 11.5 Characteristics of Monopoly
- 11.6 Profit Maximizing Output of a Monopoly Firm
- 11.7 Welfare: Perfect Competition vs Monopoly
- 11.8 Implications of Perfect Competition and Monopoly for Managerial Decision Making
- 11.9 Summary
- 11.10 Key Words
- 11.11 Self-Assessment Questions
- 11.12 Further Readings

11.1 INTRODUCTION

In the preceding unit, you have been introduced to the concept of market structure and the impact it has on the competitive behaviour of firms. You must have noted that the number and size of the firms is an important determinant of the structure of the industry and/or market.

In this unit, we shall analyze the behaviour of a firm under two different market structures, namely, pure/perfect competition and monopoly. The crucial parameter is the size of the constituent firms in relation to the total

industry's output. Throughout this unit, we go by the assumption that the firms are guided by profit maximization.

11.2 CHARACTERISTICS OF PERFECT COMPETITION

Perfect competition is a form of market in which there are a large number of buyers and sellers competing with each other in the purchase and sale of goods, respectively and no individual buyer or seller has any influence over the price. Thus perfect competition is an ideal form of market structure in which there is the greatest degree of competition.

A perfectly competitive market has the following **characteristics**:

1. There are a large number of independent, relatively small sellers and buyers as compared to the market as a whole. That is why none of them is capable of influencing the market price. Further, buyers/sellers should not have any kind of association or union to arrive at an understanding with regard to market demand/price or sales.
2. The products sold by different sellers are homogenous and identical. There should not be any differentiation of products by sellers by way of quality, variety, colour, design, packaging or other selling conditions of the product. That is, from the point of view of buyers, the products of competing sellers are completely substitutable.
3. There is absolutely no restriction on entry of new firms into the industry and the existing firms are free to leave the industry. This ensures that even in the long run the number of firms would continue to remain large and the relative share of each firm would continue to remain insignificant.
4. Both buyers and sellers in the market have perfect knowledge about the conditions in which they are operating. Buyers know the prices being charged by different competing sellers and sellers know the prices that different buyers are offering.
5. The distance between the locations of competing sellers is not significant and therefore the price of the product is not affected by the cost of transportation of goods. Buyers do not have to incur noticeable transport costs if they want to switch over from one seller to another.

The characteristics of perfect competition are summarized in Table 11.1

Number and size of distribution of sellers.	Many small sellers. No individual seller is able to exercise a significant influence over price.
Number and size distribution of buyers	Many small buyer No buyer is able to exert a significant influence over price.
Product differentiation.	No product differentiation. Decisions to buy are made on the basis of price.
Conditions of entry and exit.	Easy entry and exit. Resources are easily transferable among industries.

As mentioned in the previous unit, it is difficult to find a market that satisfies all the text book conditions of perfect competition. There are markets that come close to fulfilling these stringent conditions, but none that completely is in synchronization with all of them. You might well ask the rationale for studying this market structure if it does not exist in the real world. The answer is that perfect competition is the ideal market, and serves as a benchmark. We can use the outcomes of other markets to compare with outcomes that would have been achieved under perfect competition. For instance, if the market is competitive, prices would be lower and closer to costs, while if the market is monopolized then prices are likely to be higher. Another useful comparison relates to the concept of consumer's surplus.

Intuitively, consumer's surplus can be thought of as the difference between the maximum amount the consumer is willing to pay for a product and the amount he actually pays. Think about your purchase of a big ticket item such as a camera. You have a price in mind that is the maximum you are willing to pay. The difference between this and the price actually paid is the consumer's surplus¹.

In perfectly competitive markets, consumer's surplus is the maximum, while in monopoly markets it is low. In fact, it is the endeavour of monopolies to capture as much of the consumer's surplus as possible. When a perfectly competitive industry gets monopolized there is a transfer of surplus from the consumer to the producer. Or stated differently, the producer is able to increase his surplus (or profit) at the expense of the consumer. On the other hand, when a monopolized industry becomes competitive, there is transfer from producers to the consumers; i.e. consumers become better off when there is increased competition. An illustration of this can be gauged from the conduct of the automobile industry in India since it was deregulated in 1991. The consumers have benefited from competition in the sector and one can definitely assert that producer margins (or surplus) have declined to the benefit of the consumers.

¹ Note that you will never pay more than maximum amount.

Activity 1

Grocery stores in a large city appear to have a perfectly competitive market structure as there are many sellers and each seller is relatively small selling similar products.

- a) Do you think that grocery stores can be an example of perfect competition? Discuss.
.....
.....
.....
.....
.....
- b) What in your opinion is the market structure of grocery stores and why?
.....
.....
.....
.....
.....

11.3 PROFIT-MAXIMISING OUTPUT IN THE SHORT RUN

Having examined the rationale for studying perfectly competitive markets, let us analyze the profit-maximizing output of a profitable competitive firm in the short run. As you already know, the short run is defined as a period of time in which at least one input is fixed. Often the firm’s capital stock is viewed as the fixed input. Accordingly, this analysis assumes that the number of production facilities in the industry and the size of each facility do not change because the period being considered is too short to allow firms to enter or leave the industry or to make any changes in their operations.

Under perfect competition, since an individual firm cannot influence the market price by raising or lowering its output, the firm faces a horizontal demand curve, that is, the demand curve of any single firm is perfectly elastic – its elasticity is equal to infinity at all levels of output. If a firm charges a price slightly higher than the prevailing market price, demand for that firm will fall to zero because there are many other sellers selling exactly the same product. On the other hand, if a firm reduces its price slightly, its demand will increase to infinity and thus other firms will match the low price.

A firm under perfect competition is a price-taker and not a price-maker. Because an individual firm's demand or Average Revenue (AR) curve is horizontal under perfect competition, the Marginal Revenue (MR) curve of the firm is also horizontal coincides with the AR curve. In other words, AR and MR are constant and equal at all levels of output. You should satisfy yourself that if price (i.e. average revenue) is constant, marginal revenue will be equal to price².

The price-output determination and equilibrium of the firm under perfect competition may be explained through a numerical example. Suppose the demand and supply conditions of a product are represented by the following equations:

Aggregate Demand: $Q = 25 - 0.5 P$

Aggregate Supply: $Q = 10 + 1.0 P$

The equilibrium price would be at a point where aggregate demand equals aggregate supply

$$25 - 0.5 P = 10 + 1.0 P$$

$$\text{or } P = 10$$

Industry output at $P = 10$ is obtained by substituting this price into either the demand or supply function:

$$\begin{aligned} Q &= 10 + 1.0 (10) \\ &= 20 \end{aligned}$$

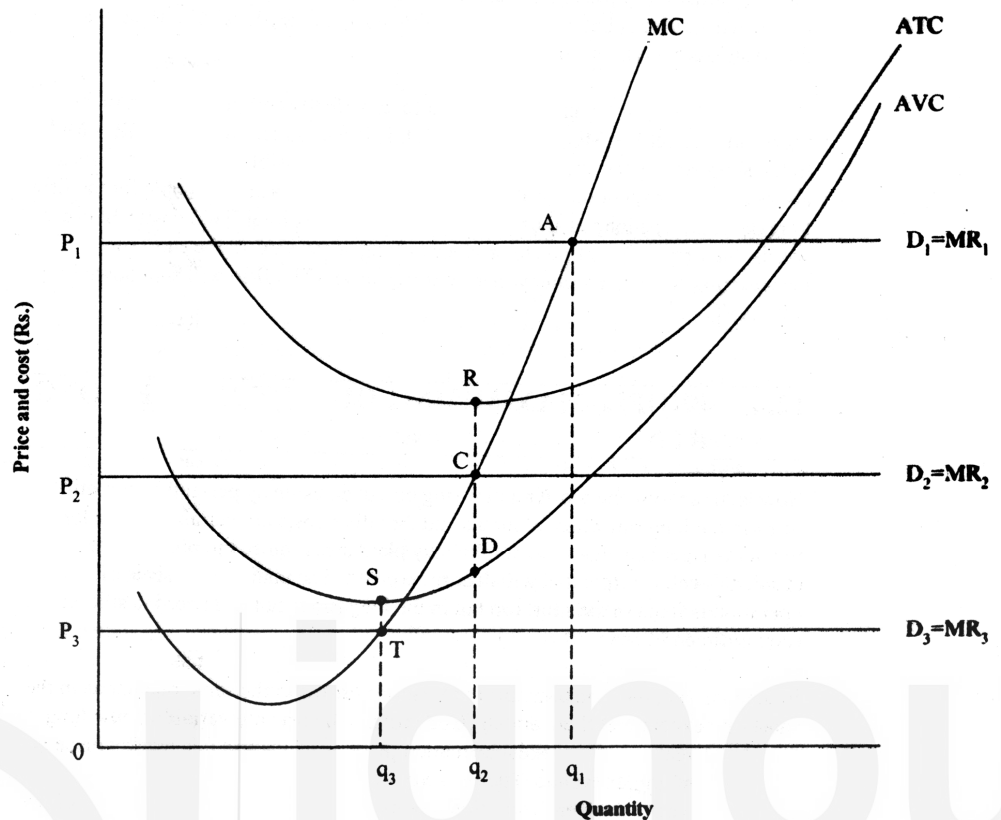
Therefore equilibrium price, $P = 10$ and equilibrium output, $Q = 20$.

Figure 11.1 shows that when the market price is at P_1 , demand and marginal revenue facing the firm are D_1 and MR_1 . The optimal output for the firm to produce is at point A, where Marginal Cost (MC) = P_1 , and the firm will produce Q_1 units of output. At Q_1 level of output, the Average Total Cost (ATC) is less than the price and the firm makes an economic profit.

Suppose the market price falls to P_2 , price equals MC at point C. Because at this level of output (Q_2) average total cost is greater than price, total cost is greater than total revenue, and the firm suffers losses. The amount of loss is the loss per unit (CR) times the number of units produced (Q_2).

² If $p=10$ and $q=1$, $TR = 10$; if $p = 10$ and $q=2$, $TR = 20$; MR is thus 10 and so on. MR will always = 10 and therefore will be the same as price as long as price is constant.

Figure 11.1: Profit Maximizing Equilibrium in the Short Run



At price level P_2 , demand is $D_2 = MR_2$, there is no way that the firm can earn a profit. This is because at every output level average total cost exceeds price ($ATC > P$). The firm will continue to produce only if it loses less by producing than by closing its operations entirely. When the firm produced zero output, total revenue would also be zero and the total cost would be the total fixed cost. The loss would thus be equal to total fixed cost. If the firm produces at $MC = MR_2$ (point C), total revenue is greater than total variable cost, because $P_2 > AVC$ at Q_2 units of output. The firm will be in a position to cover all its variable costs and still has CD times the number of units produced (Q_2) left over to pay part of its fixed cost. This way the firm suffers a smaller loss when it continues production than it shut down its operations.

At market price P_3 , demand is given by $D_3 = MR_3$. The equilibrium output Q_3 would be at T where $MC = P_3$. At this output level, since the average variable cost of production exceeds price, the firm not only loses all its fixed costs but would also lose Rs. ST per unit on its variable costs as well. The firm could improve its earnings situation by producing zero output and losing only fixed costs. In other words, when price is below average variable cost at every level of output, the short-run loss-minimizing output is zero.

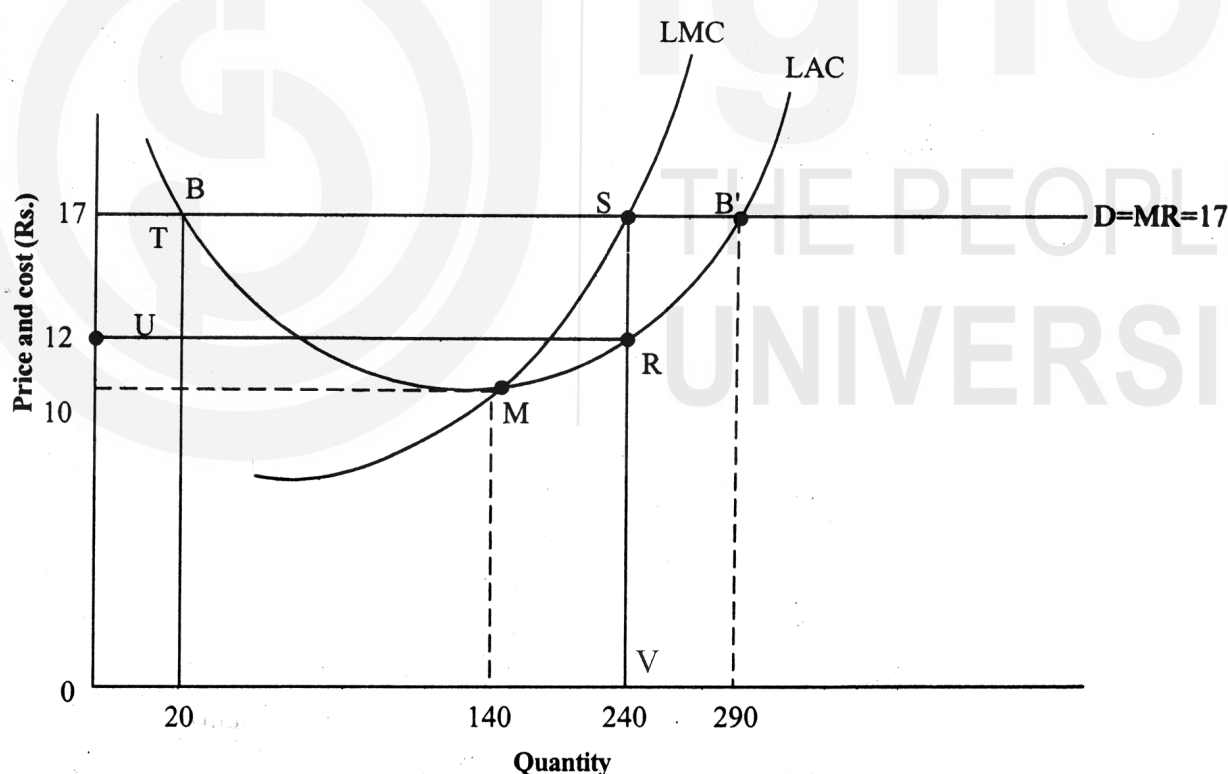
To reiterate, the profit maximizing output for a perfectly competitive firm in the short run is to set $P = MC$. Since $P = MR$, this is equivalent to setting $MR = MC$. In the short run, as the above discussion shows, it is possible for the firm to make above normal or economic profit. On the other hand, it is also possible for the firm to make losses, as long as those losses are less than its total fixed costs. In other words, the firm will continue to produce as long as $P > AVC$ in the short run, because this is a better strategy than shutting down. The firm will shut down only if $P < AVC$.

11.4 PROFIT-MAXIMISING OUTPUT IN THE LONG RUN

Now let us analyze the profit maximizing output decision by perfectly competitive firms in the long run when all inputs and therefore costs are variable. In the long run, a manager can choose to employ any plant size required to produce the efficient level of output that will maximize profit. The plant size or scale of operation is fixed in the short run but in the long run it can be altered to suit the economic conditions.

In the long run, the firm attempts to maximize profits in the same manner as in the short run, except that there are no fixed costs. All costs are variable in the long run. Here again the firm takes the market price as given and this market price is the firm's marginal revenue. The firm would increase output as long as the marginal revenue from each additional unit is greater than the marginal cost of that unit. It would decrease output when marginal cost exceeds marginal revenue. This way the firm maximizes profit by equating marginal cost and marginal revenue ($MR = MC$; as discussed above).

Figure 11.2: Profit Maximizing Equilibrium in the Long Run



The firm's long run average cost (LAC) and marginal cost (LMC) curves are shown in Figure 11.2. The firm faces a perfectly elastic demand indicating the equilibrium price (Rs. 17) which is the same as marginal revenue (i.e., $D = MR = P$). You may observe that as long as price is greater than LAC, the firm can make a profit. Therefore, any output ranging from 20 – 290 units yields some economic profit to the firm. In figure 11.2, B and B¹ are the

breakeven points, at which price equals LAC, economic profit is zero, and the firm can earn only a normal profit.

The firm, however, earns the maximum profit at output level 240 units (point S). At this point marginal revenue equals LMC and the firm would ideally select the plant size to produce 240 units of output. Note that in this situation the firm would not produce 140 units of output at point M, which is the minimum point of LAC. At this point marginal revenue exceeds marginal cost, so the firm can gain by producing more output. As shown in figure 11.2, at point S total revenue (price times quantity) at 240 units of output is equal to Rs. 4080 ($= \text{Rs. } 17 * 240$), which is the area of the rectangle OTSV. The total cost (average cost times quantity) is equal to Rs. 2,880 ($= \text{Rs. } 12 * 240$) which is the area of the rectangle OURV. The total profit is Rs. 1,200 $= (\text{Rs. } 17 - \text{Rs. } 12) * 240$, which is the area of the rectangle UTSR.

Thus, the firm would operate at a scale such that long run marginal cost equals price. This would be the most profitable situation for an individual firm (illustrated in figure 11.2). Therefore, if the price is Rs. 17.00 per unit, the firm will produce 240 units of output, generating a profit of Rs. 1,200.00. This profit is variously known as above normal, super normal or economic profit. The crucial question that one needs to ask is whether this is a sustainable situation in a perfectly competitive market i.e. whether a firm in a perfectly competitive industry can continue to make positive economic profits even in the long run? The answer is unambiguously no.

This result derives from the assumption that in a perfectly competitive market there are no barriers to entry. Recall that in a market economy, profit is a signal that guides investment and therefore resource allocation decisions. In this case, the situation will change with other prospective entrants in the industry. The economic force that attracts new firms to enter into or drives out of an industry is the existence of economic profits or economic losses respectively. Economic profits attract new firms into the industry whose entry increases industry supply. As a result, the prices would fall and the firms in the industry adjust their output levels in order to remain at profit maximization level. This process continues until all economic profits are eliminated. There is no longer any attraction for new firms to enter since they can only earn normal profits. By observing figure 11.2 you should try to work out the price that will prevail in this market in the long run when all firms are earning normal profit.

Analogous to economic profit serves as a signal to attract investment, economic losses drive some existing firms out of the industry. The industry supply declines due to exit of these firms which pushes the market prices up. As the prices have risen, all the firms in the industry adjust their output levels in order to remain at a profit maximization level. Firms continue to exit until economic losses are eliminated and economic profit becomes zero, that is, firms earn only a normal rate of profit.

Activity 2

Assume that all the assumptions of perfect competition hold true.

- a) What would be the effect of technological change in the long-run under perfect competition?

.....

.....

.....

.....

.....

- b) What conditions, in your opinion, would encourage research and development activities in the industry operating under perfect competition?

.....

.....

.....

.....

.....

11.5 CHARACTERISTICS OF MONOPOLY

Monopoly can be described as a market situation where a **single firm** controls the entire supply of a product which has **no close substitutes**. The market structure characteristics of monopoly are listed below:

- | | |
|--|-------------------------------|
| • Number and size of distribution of sellers | Single seller |
| • Number and size of distribution of buyers | Unspecified |
| • Product differentiation | No close substitutes |
| • Conditions of entry and exit | Prohibited or difficult entry |

Though perfect competition and monopoly are the two extreme cases of market structure, they both have one thing in common – they do not have to compete with other individual participants in the market. Sellers in perfect competition are so small that they can ignore each other. At the other extreme, the monopolist is the only seller in the market and has no competitors. The market or industry demand curve and that of the individual firm are the same under monopoly since the industry consists of only one firm.

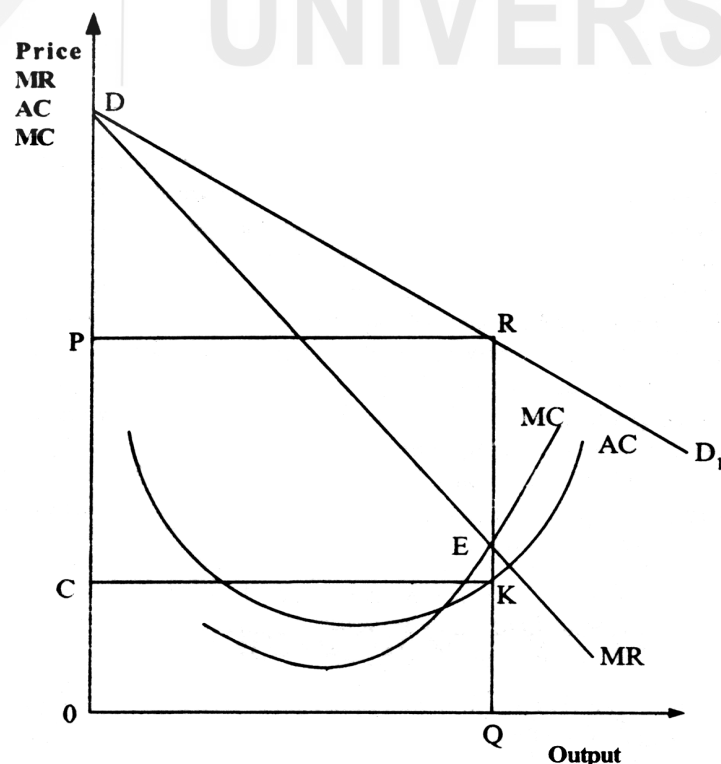
Managers of firms in a perfectly competitive market facing a horizontal demand curve would have no control over the price and they simply choose the profit maximizing output. However, the monopoly firm, facing a downward-sloping demand curve (see Figure 11.3) has power to control the

price of its product. If the demand for the product remains unchanged, the monopoly firm can raise the price as much as it wishes by reducing its output. On the other hand, if the monopoly firm wishes to sell a larger quantity of its product it must lower the price because total supply in the market will increase to the extent that its output increases. While an individual firm under perfect competition is a price-taker, a monopolist firm is a price-maker. It may, however, be noted that to have price setting power a monopoly must not only be the sole seller of the product but also sell a product which does not have close substitutes.

11.6 PROFIT MAXIMISING OUTPUT OF A MONOPOLY FIRM

Often students are tempted into thinking that since a monopolist is the only producer in the market, he will be able to charge any price for the product. While a monopolist will certainly charge a high price, it must also ensure that it is maximizing profit. Our earlier discussion proves that a profit maximizing monopoly firm determines its output at that level where its marginal cost (MC) curve intersects its downward sloping marginal revenue (MR) from below. Since the MR curve of the monopoly firm is below its average revenue or demand curve at all levels of output, and at the equilibrium output level marginal revenue is equal to marginal cost, the profit maximizing monopoly price is greater than marginal cost. You may recall, the profit maximizing price under perfect competition is equal to marginal cost. Since the demand curve of the monopoly firm is above the firm's average cost curve, the price at equilibrium output is also greater than average cost. Therefore, super-normal profits are a distinguishing feature of equilibrium under

Figure 11.3: Equilibrium output and price under monopoly



The firm would enjoy such super normal profits even in the long run because it is very difficult for new firms to enter in a monopolized market.

The determination of profit maximizing equilibrium output and price under monopoly is shown in figure 11.3. DD and MR are the downward sloping demand (or average revenue curve) and marginal revenue curves respectively of the monopoly firm. AC and MC are its average cost and marginal cost curves. At point E, MC intersects MR from below. Corresponding to E, the profit maximizing equilibrium output is OQ. At OQ output, the price is OP = QR; and average cost is OC = QK. The monopoly profits are equal to price minus average cost multiplied by output i.e., $(OP - OC) * OQ = PC * CK = PCKR$. The rectangle area PCKR represents the super normal profits of the monopoly firm.

Monopoly Power

The above analysis shows that whereas under perfect competition, price is equal to marginal cost and profits are normal in the long run; under monopoly, price is greater than marginal cost and profits are above normal even in the long run.

Therefore, the monopolist has power to charge a price which is higher than marginal cost and earn super normal profits. The extent of monopoly power of a firm can be calculated in terms of how much price is greater than marginal cost. Recall that a perfectly competitive firm sets $P = MC$. Thus the greater the difference, the greater is the monopoly power. Economist A.P. Lerner devised such an index to measure the degree of monopoly power and which has come to be known as the Lerner index. According to this index, the monopoly power of a firm is —

$$\mu = (P - MC)/P$$

Where

P = Price of the firm's product

MC = Firm's marginal cost

We know that at equilibrium output $MC = MR$ and $MR = P(1 - 1/e)$ where e is the price elasticity of demand.

$$\mu = (P - MC)/P$$

$$\mu = (P - MR)/P = 1 - (MR/P)$$

$$\text{But } (MR/P) = (1 - 1/e)$$

$$\mu = 1 - (1 - 1/e)$$

$$\mu = 1/e$$

The monopoly power of a firm is inversely related to elasticity of demand for its product. The less elastic the demand for its product, the greater would be its monopoly power, and *vice versa*. As we have discussed in Block 2, elasticity of demand depends on the number and closeness of the substitutes

available for a product. In the real world we find some essential goods and services like life saving medicines, petroleum, cooking gas, railways etc. enjoy a high degree monopoly power because the demand for these products is highly inelastic. Left to itself the monopoly could price such inelastic products at rates that do not meet the social objectives of the government and policy makers. Thus we often witness government intervention in monopolies. For example, Railway ticket prices are fixed by the government and electricity tariffs are set by a regulatory authority. The reason why monopolies need to be regulated is discussed in the next section.

11.7 WELFARE: PERFECT COMPETITION VS MONOPOLY

Our discussion reveals that in a pure monopoly price will generally be greater than marginal cost and that the firm is able to generate super normal profits at the expense of welfare of consumers and society even in the long run. Recall that key conditions that give rise to monopolies are economies of scale and barriers to entry. On the other hand, production processes like food processing, textiles, garments, wood and furniture, it is relatively easy to enter the market as a supplier – for example, capital requirements are low and sunk costs are also low. Many service industries like travel agencies fall into this category. In such industries, competition ensures that prices are set ‘right’ and moreover the threat of entry ensures that prices never exceed long-run average cost (for example, marginal companies in the industry cannot persistently earn above average profits). Moreover, competition also ensures that price equals long-run marginal cost. Hence the price of a good accurately reflects the opportunity cost of manufacturing it.

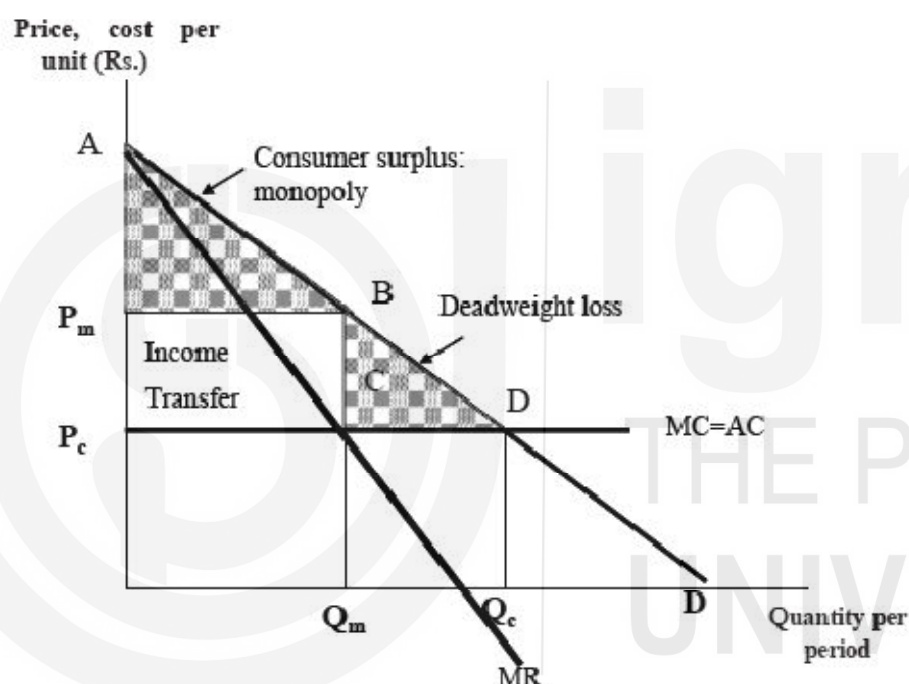
Problems arise from leaving everything to the market, however when a situation of monopoly occurs. In economists’ jargon, there are economies of scale to be exploited when one company meets market demand. There are typically also major barriers to entry in such industries. Most public utilities – electricity generation, water supply, gas supply and perhaps national telecommunications systems – have technologies of this sort. There are several special problems for these industries.

First, their size and capital intensity often puts particular strain on private capital markets in satisfying their investment needs. In India, in the 1990s strain was felt instead on the public coffers, and this was a major factor behind the move towards disinvestment and privatization. Hence, while for example automobile or chemicals manufacture are also characterized by huge scale economies, governments have rarely seen it as their role to regulate companies in these industries. The question for policy makers is what to do about natural monopolies like power and water supply. Left to themselves, they will charge monopoly prices and restrict output. The absence of any competitive threat will also probably leave such organizations wasteful, inefficient and sluggish. Since all costs can be passed on to the consumers,

there will be little incentive for managers to keep them under control. Experience from, for example, the railways suggests that it will not be long before the absence of competitive pressures may damage the motives for innovation and change, so crucial in such capital-intensive sectors. Thus in some cases a regulator is appointed who must fix the natural monopolist's price. In India, privatization of power and telecommunications has been accompanied by the creation of a regulator, while there is no such institution for cement, automobile or chemical industry.

The above discussion can also be illustrated with the help of Figure 11.4. Assume a perfectly competitive industry. We know that price would be P_c and quantity supplied Q_c .

Figure 11.4: Evaluation of Monopoly



The consumer's surplus will be the area $P_c AD$. Now consider output and price of the profit maximizing monopolist. As indicated in the figure, price would be P_m and quantity would be Q_m . Notice that the monopolist will charge a higher price and produce a lower quantity as expected. The consumer surplus is reduced to $P_m AB$. The rectangle $P_c P_m BC$ that was part of consumer surplus under competition is now economic profit for the monopolist. This economic profit represents income redistribution from consumers to producers. Further, there is also a deadweight loss to society represented by the area BCD that represents loss of consumer surplus that accrued under competition, but is lost to society because of lower production levels under monopoly.

If we now consider the reverse case i.e. a monopoly being broken to foster competition, the result will be transfer of income from producers to consumers and elimination of deadweight loss. Herein lies the economic

basis for regulation of monopoly firms. It is to generate the outcomes of competitive markets and pass these benefits to consumers in the form of lower prices. If competition exists in markets then arguably, that is the best regulation. If it does not, and the industry is envisaged to play a social role, regulation of monopoly becomes an important policy objective.

Activity 3

Give few examples of market situation where monopoly exists and explain.

.....

.....

.....

.....

.....

11.8 IMPLICATIONS OF PERFECT COMPETITION AND MONOPOLY FOR MANAGERIAL DECISION MAKING

The assumptions underlying perfect competition market are very restrictive. Few markets are found with characteristics of many small sellers, easy entry and exit, and an undifferentiated product. Normally, a majority of modern industries operate under conditions of oligopoly or monopolistic competition. You will study these two market structures in detail in Unit 13.

Perfect competition and monopoly are the two extreme market conditions which we rarely come across in the real world of business. Then the question arises as to why study them? It is useful to think of perfect competition and pure monopoly as extremes with other market structures placed in between. There are many industries that have most of the characteristics of perfect competition or monopoly. The two extreme models therefore serve as benchmarks and provide guidance in making decisions.

Consider the following case. A multinational soft drink company P was in bankruptcy for the second time. The President of the Company tried to sell it to soft drink Company C, but the C Company wanted no part of the deal. In order to reduce costs the President purchased a large supply of recycled 12-ounce beer bottles. At that time, soft drink P and C were sold in six ounce bottles. Initially Company P priced the bottles at 10 cents, twice the amount of the original six ounce bottles, but with little success. Then the President of the Company had the brilliant idea of selling the 12 ounce bottles of softdrink P at the same price as the six ounce bottles of softdrink C. Then scale took off and shortly softdrink P was out of bankruptcy and soon making a very nice profit.

Softdrink P pricing decision was clearly crucial to the life of the firm. The primary background necessary for understanding the pricing decision is a good understanding of the law of demand – i.e. as price goes up, demand goes down – and some understanding of the amount by which a price increase effects a quantity decrease – i.e. the price elasticity of demand. We will start by examining the polar cases of pricing under perfect competition and pricing under monopoly, and then move on to examining Softdrink P and C company's situation.

Alfred Marshall, a famous 19th Century economist, used a fish market as an example of perfect competition. For the sake of argument, consider a fish seller selling cod. How would s/he price his product? First, s/he would look around and find out at what price her/his numerous competitors were selling cod. s/he certainly could not price above the competitors; since cod is pretty much identical and consumers should not care from whom they purchase. Furthermore, in fish markets, it is quite easy for consumers to compare prices. So, if s/he priced above her/his competitors, s/he would not sell any fish. Suppose s/he decided to price below her/his competitors. All of the customers would certainly purchase from him. However, if s/he were still making a profit, the other competitors would also be making a profit at the lower price and would match the price cut in order to retain their customers. They may even consider lowering price more, if they could still make a profit and capture further customers.

This reasoning, along with the ease of entry for new fish seller, if there is a profit to be made (which prevents collusion among fish seller already in the market), ensures that the price being charged is equal to the cost of supplying an additional fish, or the marginal cost. A fish seller will be a price-taker, setting her/his price identically to her/his competitors' prices. A firm is a monopoly if it has exclusive control over the supply of a product or service. Therefore, a monopolist, in her/his pricing decisions, cannot consider the pricing decision of rival firms. So, what does s/he consider?

The smart monopolist considers the incremental effect of his decision, i.e. what is the revenue to be received from selling one additional unit of a product and what are the costs of selling one additional unit of a product. Certainly, if the costs of selling one additional unit of a product exceed the revenues, the monopolist would certainly not want to sell that additional product. The law of demand says that s/he could raise the price of her/his product and thus sell less. Alternatively, if the revenues of selling an additional unit of a product exceed the costs of selling that unit, the monopolist should want to sell more units. The law of demand says that s/he could sell more by lowering his price.

Thus, by setting the price correctly, the monopolist can sell the exact number of units such that the costs of selling one additional unit exactly equals the revenues of selling the additional unit, which, by the above reasoning, is the only optimal price. However, there is an additional complication: the costs of

selling one additional unit do not include any part of the salary of the CEO or the rental costs of the plant, both must be paid whether or not the additional unit is sold. Thus, in the long run, if a monopolist cannot cover his overhead by pricing in the optimal manner, he should shut down.

The situation involving soft drink P and soft drink C clearly differs from either of the above scenarios, but what can we learn from both the cases? First, P clearly saw that C was pricing the six-ounce bottles at 5 cents. By pricing the 12-ounce bottles at 5 cent also, P made the bet that C would not cut its price. C did not see the need to cut price because its product was different from P and it did not fear losing many of its customers. Whether the gain in revenues resulting from increased demand would offset the loss in revenue from the lower price depends on the price elasticity of demand. The price elasticity of demand faced by soft drink P company depends on soft drink C response to the price cut and the consumer's responses. As we saw above, P made the assumption that C would not cut price. P counted on a highly elastic consumer response, that is the percentage change in quantity purchased by the consumer due to the lower price, and therefore profits would accrue to P.

What other concerns you think played a part in the soft drink P company's decision?

11.9 SUMMARY

In this unit, you have studied the market forces operating in perfect competition and pure monopoly; and the pricing and output decisions in these two market structures. The perfect competition model assumes a large number of small sellers and buyers, identical products, and an easy entry and exit conditions. In perfect competition, firms face a horizontal demand curve at equilibrium price. Price is determined by the interaction of the market supply and demand curves. Since no single firm has control over price, the objective of managers is to determine the level of output that maximizes profit.

The perfectly competitive firm maximizes profit at a point where price equals marginal cost. The firm can make an economic profit or loss in the short run, depending on market price. If the price drops below average variable cost, the firm should shut down. Or even if the firm is making a profit in the short run, it may wish to change its plant size or capacity in the long run in order to earn more profit.

The monopolist is a single seller of a differentiated product. Entry into the market is difficult or prohibited. Being the single seller, the monopolist has power over price. For maximizing profits, the firm produces until marginal revenue equals marginal cost. This way the monopolist earns economic profits in both the short run and long run as well because entry is restricted for new firms.

In the real world, few market structures meet the restrictive assumptions for perfect competition or monopoly. Still, these two models are useful because many industries have the characteristics of perfect competition or monopoly. Moreover, the perfectly competitive model serves as a benchmark for evaluating the performance of actual markets and provides guidance for public policy.

11.10 KEY WORDS

Economic Costs include normal profits.

Economic Profit represents an above-normal profit situation.

Equilibrium of a Firm ($MR = MC$) represents profit maximizing price-output combination. In a situation where maximum profits mean a loss, the equation gives loss.

Equilibrium of an Industry is stated in terms of the condition of normal profit $AR = AC$ such that the size and structure of the industry in terms of number of firms are strictly defined.

Differentiated Products: Products which are similar in nature but differ in terms of packing, look etc.

Profit Maximization: It is the condition where marginal revenue and marginal cost are in equilibrium.

11.11 SELF-ASSESSMENT QUESTIONS

1. Vegetable market is an example, closest to the pure competition. Discuss.
2. Suppose a firm A has
$$Q = 20 - 0.3 P$$
AndAggregate supply as
$$Q = 10 + 0.2 P$$
What would be the equilibrium price and the equilibrium output of the firm A?
3. Suppose a small locality has a single grocery store selling multiple products.
 - a. Is it a monopoly?
 - b. If yes, then give arguments in support of your answer.
4. Discuss the relevance of perfect competition and monopoly in the present context.

11.12 FURTHER READINGS

Mote, V. L., Paul, S., & Gupta, G. S. (2016). *Managerial Economics: Concepts and Cases*, Tata McGraw-Hill, New Delhi.

Maurice, S. C., Smithson, C. W., & Thomas, C. R. (2001). *Managerial Economics: Applied Microeconomics for Decision Making*. McGraw-Hill Publishing.

Dholakia, R., & Oza, A. N. (1996). *Microeconomics for Management Students*. Oxford University Press, Delhi.



ignou
THE PEOPLE'S
UNIVERSITY

UNIT 12 PRICING UNDER MONOPOLISTIC AND OLIGOPOLISTIC COMPETITION

Objectives

After going through this unit, you should be able to:

- **describe** the concept of the pricing decisions under monopolistic competition in short run as well as long run;
- **explain** the concept of product differentiation with special reference to monopolistic competition;
- **differentiate** between monopolistic competition and oligopoly;
- **apply** models of oligopoly behavior to real world situations.

Structure

- 12.1 Introduction
- 12.2 Monopolistic Competition
- 12.3 Price and Output Determination in Short run
- 12.4 Price and Output Determination in Long run
- 12.5 Oligopolistic Competition
- 12.6 Summary
- 12.7 Key Words
- 12.8 Self-Assessment Questions
- 12.9 Further Readings

12.1 INTRODUCTION

Pricing decisions tend to be the most important decisions made by any firm in any kind of market structure. The concept of pricing has already been discussed in unit 11. The price is affected by the competitive structure of a market because the firm is an integral part of the market in which it operates. We have examined the two extreme markets viz. monopoly and perfect competition in the previous unit. In this unit the focus is on monopolistic competition and oligopoly, which lie in between the two extremes and are therefore more applicable to real world situations.

Monopolistic competition normally exists when the market has many sellers selling differentiated products, for example, retail trade, whereas oligopoly is said to be a stable form of a market where a few sellers operate in the market and each firm has a certain amount of share of the market and the firms

recognize their dependence on each other. The features of monopolistic and oligopoly are discussed in detail in this unit.

12.2 MONOPOLISTIC COMPETITION

Edward Chamberlin, who developed the model of monopolistic competition, observed that in a market with large number of sellers, the product so individual firms are not at all homogeneous, for example, soaps used for personal care. Each brand has a specific characteristic, be it packaging, fragrance, look etc., though the composition remains the same. This is the reason that each brand is sold individually in the market. This shows that each brand is highly differentiated in the minds of the consumers. The effectiveness of the particular brand may be attributed to continuous usage and heavy advertising.

As defined by Joe S. Bain ‘Monopolistic competition is found in the industry where there are a large number of sellers, selling differentiated but close substitute products’. Take the example of soaps L and C. Both are soaps for personal care but the brands are different. Under monopolistic competition, the firm has some freedom to fix the price i.e. because of differentiation a firm will not lose all customers when it increases its price.

Monopolistic competition is said to be the combination of perfect competition as well as monopoly because it has the features of both perfect competition and monopoly. It is closer in spirit to a perfectly competitive market, but because of product differentiation, firms have some control over price. The characteristic features of monopolistic competition are as follows:

- **A large number of sellers:** Monopolistic market has a large number of sellers of a product but each seller acts independently and has no influence on others.
- **A large number of buyers:** Just like the sellers, the market has a large number of buyers of a product and each buyer acts independently.
- **Sufficient Knowledge:** The buyers have sufficient knowledge about the product to be purchased and have a number of options available to choose from. For example, we have a number of petrol pumps in the city. Now it depends on the buyer and the ease with which s/he will get the petrol decides the location of the petrol pump. Here accessibility is likely to be an important factor. Therefore, the buyer will go to the petrol pump where s/he feels comfortable and gets the petrol filled in the vehicle easily.
- **Differentiated Products:** The monopolistic market categorically offers differentiated products, though the difference in products is marginal, for example, toothpaste.
- **Free Entry and Exit:** In monopolistic competition, entry and exit are quite easy and the buyers and sellers are free to enter and exit the market at their own will.

Nature of the Demand Curve

The demand curve of the monopolistic competition has the following characteristics:

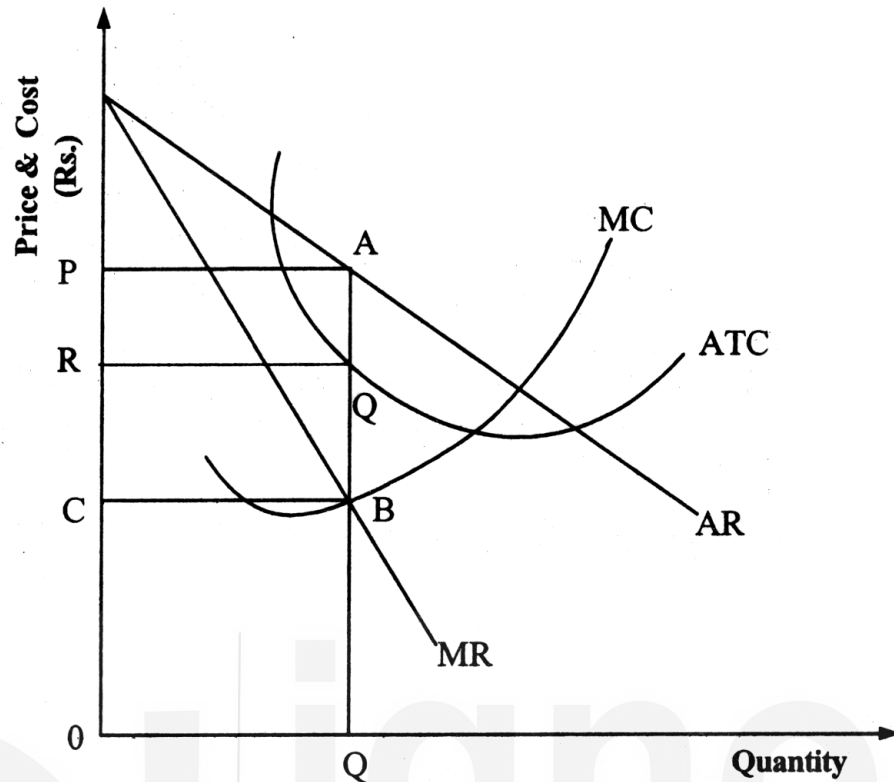
- **Less than perfectly elastic:** In monopolistic competition, no single firm dominates the industry and due to product differentiation, the product of each firm seems to be a close substitute, though not a perfect substitute for the products of the competitors. Due to this, the firm in question has high elasticity of demand.
- **Demand curve slopes downward:** In monopolistic competition, the demand curve facing the firm slopes downward due to the varied tastes and preferences of consumers attached to the products of specific sellers. This implies that the demand curve is not perfectly elastic.

12.3 PRICE AND OUTPUT DETERMINATION IN SHORT RUN

In monopolistic competition, every firm has a certain degree of monopoly power i.e. every firm can take initiative to set a price. Here, the products are similar but not identical, therefore there can never be a unique price but the prices will be in a group reflecting the consumers' tastes and preferences for differentiated products.

In this case the price of the product of the firm is determined by its cost function, demand, its objective and certain government regulations, if there are any. As the price of a particular product of a firm reduces, it attracts customers from its rival groups (as defined by Chamberlin). Say for example, if 'S company' TV reduces its price by a substantial amount or offers discount, then the customers from the rival group who have loyalty for, say 'L TV company', tend to move to buy 'S company's' TV sets.

As discussed earlier, the demand curve is highly elastic but not perfectly elastic and slopes downwards. The market has many firms selling similar products, therefore the firm's output is quite small as compared to the total quantity sold in the market and so its price and output decisions go unnoticed. Therefore, every firm acts independently and for a given demand curve, marginal revenue curve and cost curves, the firm maximizes profit or minimizes loss when marginal revenue is equal to marginal cost. Producing an output of Q selling at price P maximizes the profits of the firm.

Figure 12.1: Short run equilibrium under monopolistic competition

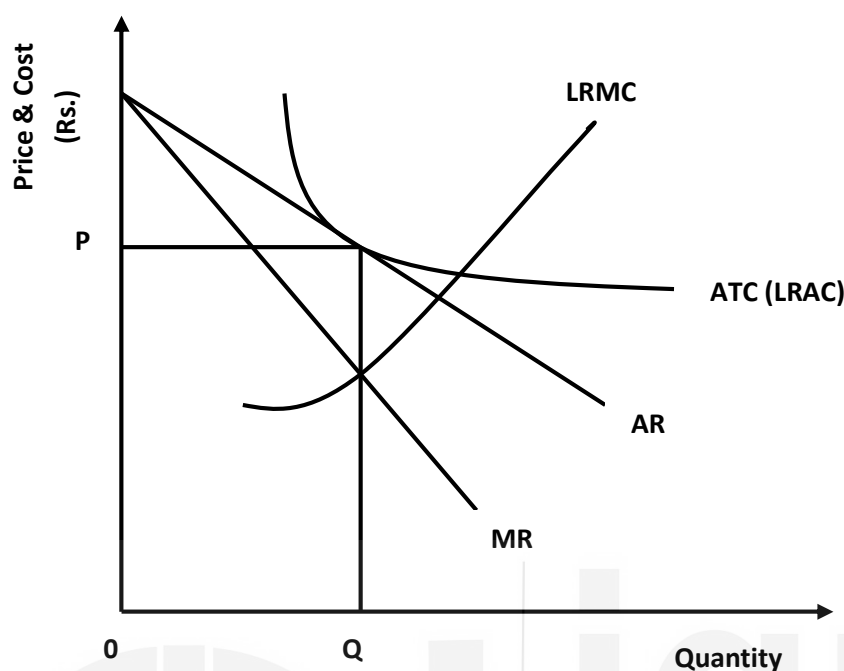
In the short run, a firm may or may not earn profits. Figure 12.1 shows the firm, which is earning economic profits. The equilibrium point for the firm is at price P and quantity Q and is denoted by point A. Here, the economic profit is given as area PAQR. The difference between this and the monopoly case is that here the barriers to entry are low or weak and therefore new firms will be attracted to enter. Fresh entry will continue to enter as long as there are profits. As soon as the super normal profit is competed away by new firms, equilibrium will be attained in the market and no new firms will be attracted in the market. This is the situation corresponding to the long run and is discussed in the next section.

12.4 PRICE AND OUTPUT DETERMINATION IN LONG RUN

We have discussed the price and output determination in the short run. We now discuss price and output determination in the long run. You will notice that the long run equilibrium decision is similar to perfect competition. The core of the discussion under this head is that economic profits are eliminated in the long run, which is the only equilibrium consistent with the assumption of low barriers to entry. This occurs at an output where price is equal to the long run average cost. The difference between monopolistic competition and perfect competition is that in monopolistic competition the point of tangency is downward sloping and does not occur at minimum of the average cost curve and this is because the demand curve is downward sloping¹.

¹ You should appreciate that $P=AC$ is the only compatible long run equilibrium under both perfect competition and monopolistic competition. The reason is that there are no entry barriers. However, because the demand curve is downward sloping in monopolistic competition the point at which $P=AC$ occurs to the left of the minimum point of the average cost curve, rather than at the minimum point, as in perfect competition

Figure 12.2: Long run equilibrium under monopolistic competition



Looking at figure 12.2, under monopolistic competition in the long run we see that LRAC is the long run average cost curve and LRMC the long run marginal curve. Let us take a hypothetical example of a firm in a typical monopolistic situation where it is making substantial amount of economic profits. Here it is assumed that the other firms in the market are also making profits. This situation would then attract new firms in the market. The new firms may not sell the same products but will sell similar products. As a result, there will be an increase in the number of close substitutes available in the market and hence the demand curve would shift downwards since each existing firm would lose market share. The entry of new firms would continue as long as there are economic profits. The demand curve will continue to shift downwards till it becomes tangent to LRAC at a given price P and output at Q as shown in the figure. At this point of equilibrium, an increase or decrease in price would lead to losses. In this case the entry of new firms would stop, as there will not be any economic profits. Due to free entry, many firms can enter the market and there may be a condition where the demand falls below LRAC and ultimately suffers losses resulting in the exit of the firms. Therefore under the monopolistic competition free entry and exit must lead to a situation where demand becomes tangent to LRAC, the price becomes equal to average cost and no economic profit is earned. It can thus be said that in the long run the profits peter out completely.

One of the interesting features of the monopolistically competitive market is the variety available due to product differentiation. Although firms in the long run do not produce at the minimum point of their average cost curve, and thus there is excess capacity available with each firm, economists have rationalized this by attributing the higher price to the variety available.

Further, consumers are willing to pay the higher price for the increased variety available in the market.

Activity 1

1. It is a usual practice for the customers to go to the market and purchase household goods like toothpastes, soaps, detergents etc. List (any five) such branded items along with their competitors having a substantial share in the market.

.....

2. 'In a monopolistic competition, the profits in the long run evade off completely'. Briefly discuss the statement taking into account the present trends.

.....

12.5 OLIGOPOLISTIC COMPETITION

We define oligopoly as the form of market organization in which there are few sellers of a homogeneous or differentiated product. If there are only two sellers, we have a duopoly. If the product is homogeneous, we have a pure oligopoly. If the product is differentiated, we have a differentiated oligopoly. While entry into an oligopolistic industry is possible, it is not easy (as evidenced by the fact that there are only a few firms in the industry).

Oligopoly is the most prevalent form of market organization in the manufacturing sector of most nations, including India. Some oligopolistic industries in India are automobiles, primary aluminum, steel, electrical equipment, glass, breakfast cereals, cigarettes, and many others. Some of these products (such as steel and aluminum) are homogeneous, while others (such as automobiles, cigarettes, breakfast cereals, and soaps and detergents) are differentiated. Oligopoly exists also when transportation costs limit the market area. For example, even though there are many cement producers in India, competition is limited to the few local producers in a particular area.

Since there are only a few firms selling a homogeneous or differentiated production oligopolistic markets, the action of each firm affects the other firms in the industry and vice versa. For example, when automobile company GM introduced price rebates in the sale of its automobiles, automobile company F and M immediately followed with price rebates of their own. Furthermore, since price competition can lead to ruinous price wars,

oligopolists usually prefer to compete on the basis of product differentiation, advertising, and service. These are referred to as non price competition. Yet, even here, if GM mounts a major advertising campaign, F and M are likely to soon respond in kind. When softdrink company P mounted a major advertising campaign in the early 1980s softdrink company C responded with a large advertising campaign of its own in the United States.

From what has been said, it is clear that the distinguishing characteristic of Oligopoly is the interdependence or rivalry among firms in the industry. This is the natural result of fewness. Since an oligopolist knows that its own actions will have a significant impact on the other oligopolists in the industry, each oligopolist must consider the possible reaction of competitors in deciding its pricing policies, the degree of product differentiation to introduce, the level of advertising to be undertaken, the amount of service to provide, etc. Since competitors can react in many different ways (depending on the nature of the industry, the type of product, etc.) We do not have a single oligopoly model but many—each based on the particular behavioural response of competitors to the actions of the first. Because of this interdependence, managerial decision making is much more complex under oligopoly than under other forms of market structure. In what follows we present some of the most important oligopoly models. We must keep in mind, however, that each model is at best incomplete.

The sources of oligopoly are generally the same as for monopoly. That is, (1) economies of scale may operate over a sufficiently large range of outputs as to leave only a few firms supplying the entire market; (2) huge capital investments and specialized inputs are usually required to enter an oligopolistic industry (say, automobiles, aluminum, steel, and similar industries), and this acts as an important natural barrier to entry; (3) a few firms may own a patent for the exclusive right to produce a commodity or to use a particular production process; (4) established firms may have a loyal following of customers based on product quality and service that new firms would find very difficult to match; (5) a few firms may own or control the entire supply of a raw material required in the production of the product; and (6) the government may give a franchise to only a few firms to operate in the market. The above are not only the sources of oligopoly but also represent the barriers to other firms entering the market in the long run. If entry were not so restricted, the industry could not remain oligopolistic in the long run. A further barrier to entry is provided by limit pricing, whereby, existing firms charge a price low enough to discourage entry into the industry. By doing so, they voluntarily sacrifice short-run profits in order to maximize long-run profits.

As discussed earlier oligopolies can be classified on the basis of type of product produced. They can be homogeneous or differentiated. Steel, Aluminium etc. come under homogeneous oligopoly and television, automobiles etc. come under heterogeneous oligopoly. The type of product produced may affect the strategic behaviour of oligopolists. According to

economists, two contrasting behaviour of oligopolists arise that is the **cooperative oligopolists** where an oligopolist follows the pattern followed by rival firms and the **non-cooperative oligopolists** where the firm does not follow the pattern followed by rival firms. For example, a firm raises price of its product, the other firms may keep their prices low so as to attract the sales away from the firm, which has raised its price. But as stated above, price is not the only factor of competition. As a matter of fact other factors on the basis of which the firms compete include advertising, product quality and other marketing strategies. Therefore, we normally have four general oligopolistic market structures, two each under cooperative as well as non-cooperative structures. We have firms producing homogeneous and differentiated products under each of the two basic structures. All these differences exist in the oligopolistic market. This shows that each firm tries to make an impact in the existing market structure and have an effect on the rival firms. This tends to be a distinguishing characteristic of an oligopolistic market.

Activity 2

1. List five products along with the names of the companies following homogeneous oligopolistic competition.

.....

.....

.....

.....

2. List five products along with the names of the companies following heterogeneous oligopolistic competition.

.....

.....

.....

.....

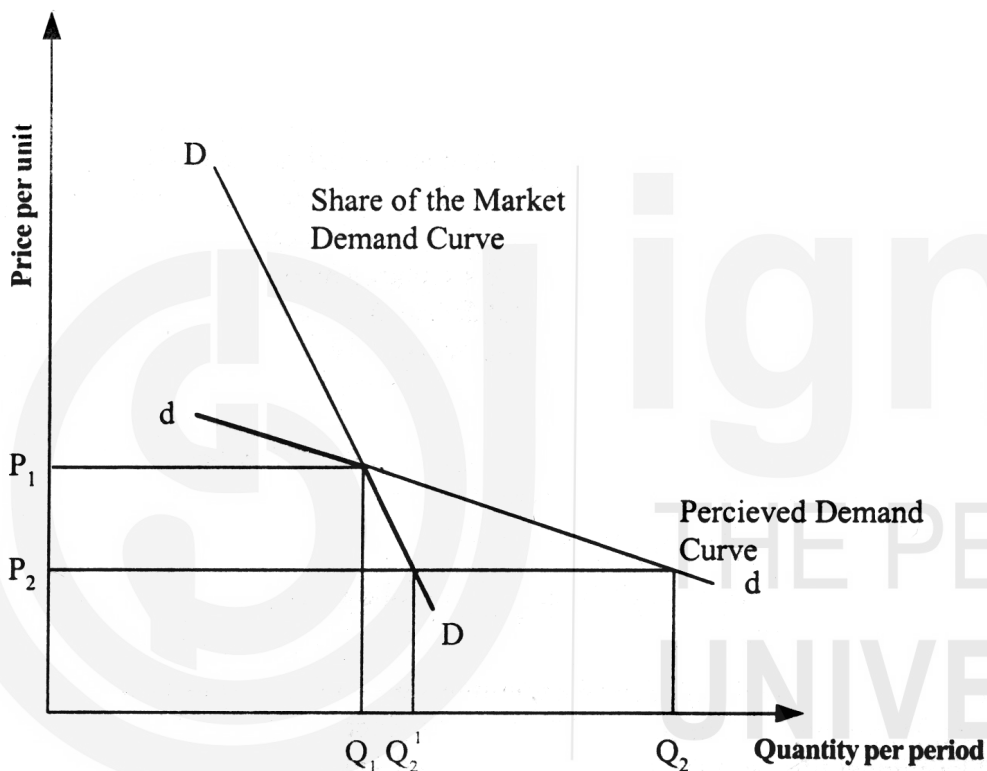
Price Rigidity: Kinked Demand Curve

Our study of pricing and market structure has so far suggested that a firm maximizes profit by setting $MR=MC$. While this is also true for oligopoly firms, it needs to be supplemented by other behavioural features of firm rivalry. This becomes necessary because the distinguishing feature of oligopolistic markets is interdependence. Because there are a few firms in the market, they also need to worry about rival firm's behaviour. One model explaining why oligopolists tend not to compete with each other on price, is the kinked demand curve model of Paul Sweezy. In order to explain this characteristic of price rigidity i.e. prices remaining stable to a great extent, Sweezy suggested the kinked demand curve model for the oligopolists. The kink in the demand curve arises from the asymmetric behaviour of the firms. The proponents of the hypothesis believe that competitors normally follow

price decreases i.e. they show the cooperative behaviour if a firm reduces the price of its products whereas they show the non-cooperative behaviour if a firm increases the price of its products.

Let us start from P_1 in Figure 12.3. If one firm reduces its price and the other firms in the market do not respond, the price cutter may substantially increase its sales. This result is depicted by the relative elastic demand curve, dd . For example, a price decrease from P_1 to P_2 will result in a movement along dd and increase sales from Q_1 to Q_2 as customers take advantage of the lower price and abandon other suppliers. If the price cut is matched by other firms, the increase in sales will be

Figure 12.3: Demand curve for an Oligopolist



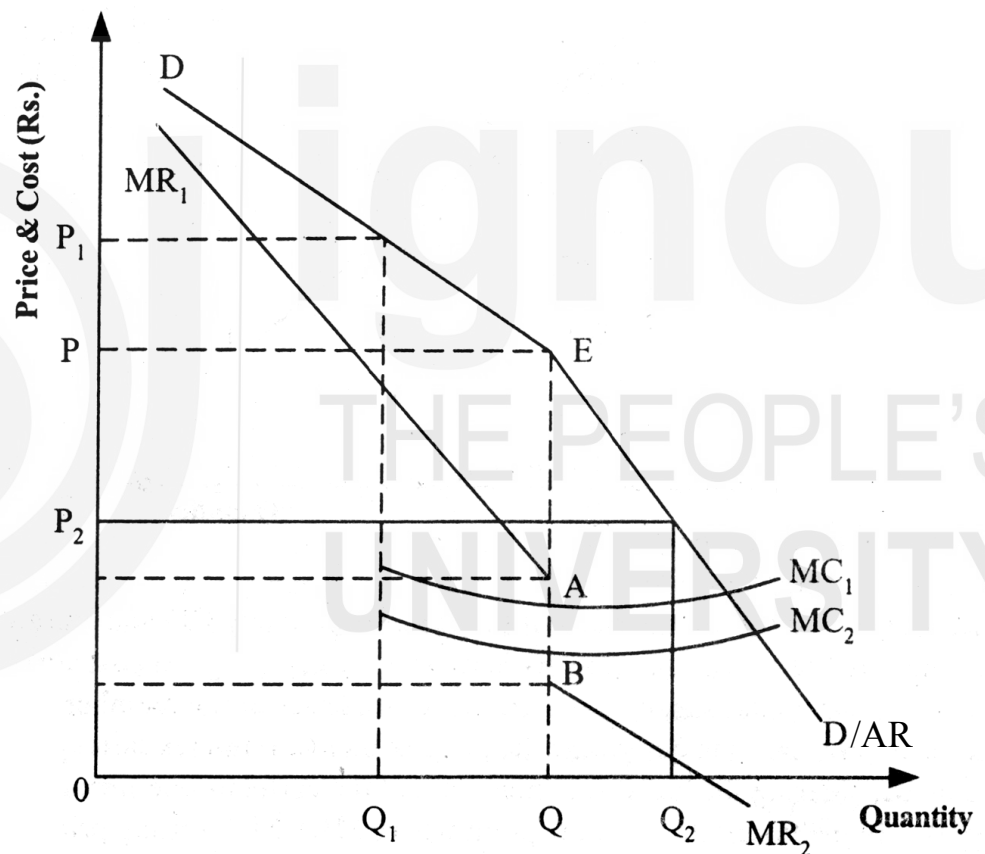
less. Since other firms are selling at the same price, any additional sales must result from increased demand for the product. Thus the effect of price reduction is a movement down the relatively inelastic demand curve, DD , then the price reduction from P_1 to P_2 only increases sales to Q_2^1 .

Here we assume that P_1 is the initial price of the firm operating in a non-cooperative oligopolistic market structure producing Q_1 units of output. P is also the point of kink in the demand curve and is the initial price and DD is the relatively elastic demand curve above the existing price P_1 . When the firm is operating in the non-cooperative oligopolistic market it results in decline in sales if it changes its price to P_1 . Now if the firm reduces its price below P_1 say P_2 , the other firms operating in the market show a cooperative behaviour and follow the firm. This is shown in the figure as the curve below the existing price P_1 . The true demand curve for the oligopolistic market is dD and has the kink at the existing price P_1 . The demand curve has two linear curves, which are joined at price P .

Associated with the kinked demand curve is a marginal revenue function. This is shown in Figure 12.4. Marginal Revenue for prices above the kink is given by MR_1 and below the kink as MR_2 . At the kink, marginal revenue has a discontinuity at AB and this depends on the elasticity of the different parts of the demand curve.

Therefore, in the presence of a kinked demand curve, firm has no motive to change its price. If the firm is a profit maximizing firm where $MR=MC$, it would not change its price even if the cost changes. This situation occurs as long as changes in MC fall within the discontinuous range i.e. AB portion. The firm following kinked model has a U-shaped marginal cost curve MC. The new MC curve will be MC_1 or MC_2 and will remain in the discontinued area and the equilibrium price remains the same at P.

Figure 12.4 Kinked demand curve



Price Competition: Cartels and Collusion

Cartel Profit Maximization

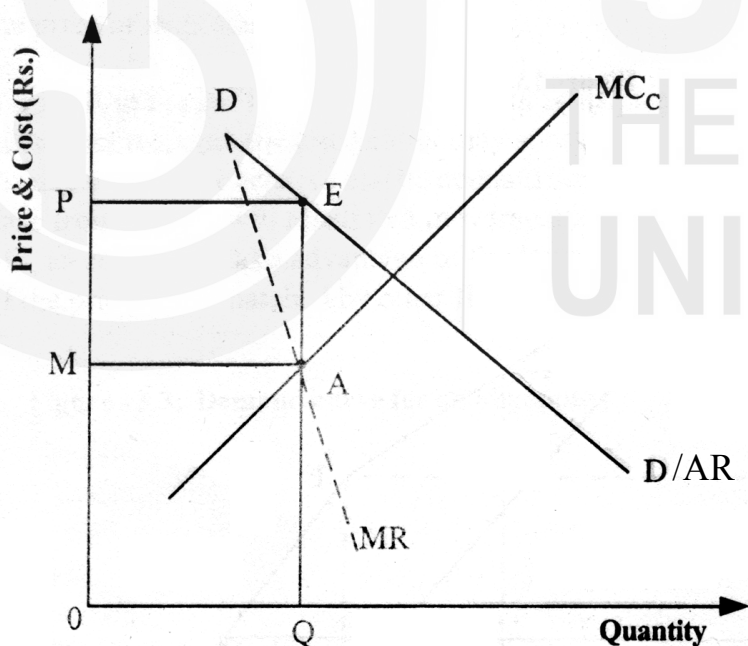
We already know now that in an oligopolistic competition, the firms can compete in many ways. Some of the ways include price, advertising, product quality, etc. Many firms may not like competition because it could be mutually disadvantageous. For example, advertising. In this case many oligopolies end up selling the products at low prices or doing high advertising resulting in high costs and making lower profits than expected. Therefore, it is possible for the firms to come to a consensus and raise the price together,

increasing the output without much reduction in sales. In some countries this kind of collusive agreement is illegal e.g. USA but in some it is legal. The most extreme form of the collusive agreement is known as a *cartel*.

A cartel is a market sharing and price fixing arrangement between groups of firms where the objective of the firm is to limit competitive forces within the market. The forms of cartels may differ. It can be an **explicit collusive agreement** where the member firms come together and may reach a consensus regarding the price and market sharing or **implicit cartel** where the collusion is secretive in nature.

Throughout the 1970s, the Organization of Petroleum Exporting Countries (OPEC) colluded to raise the price of crude oil from under \$3 per barrel in 1973 to over \$30 per barrel in 1980. The world awaited the meeting of each OPEC price-setting meeting with anxiety. By the end of 1970s, some energy experts were predicting that the price of oil would rise to over \$100 per barrel by the end of the century. Then suddenly the cartel seemed to collapse. Prices moved down, briefly touching \$10 per barrel in early 1986 before recovering to \$18 per barrel in 1987. OPEC is the standard example used in textbooks when explaining cartel behaviour. The cartel profit maximizing theory can be explained using figure 12.5.

Figure 12.5: Cartel profit maximization



The market demand for all members of the cartel is given by DD and marginal revenue (represented by dotted line) as MR. The cartels marginal cost curve given by MC_C is the horizontal sum of the marginal cost curves of the member firms. In this the basic problem is to determine the price, which maximizes cartel profit. This is done by considering the individual members of the cartel as one firm i.e. a monopoly. In the figure this is at the point where $MR = MC_C$, setting price = P. The problem is regarding the allocation of output within the member firms. Normally a quota system is quite popular,

whereby each firm produces a quantity such that its $MC = MC_C$. One serious problem that arises from this analysis is that while the joint profits of the cartel as a whole are maximized, each individual member of the cartel has an incentive to cheat on its quota. This is because the price for the product is greater than the members marginal cost of production. This implies that an individual member can increase its profit by increasing production. What would happen if all members did the same? The market sharing arrangement will breakdown and the cartel would collapse. Here lies the inherent instability of cartel type arrangement and can be summarized as follows.

There is an incentive for the cartel as a whole to restrict output and raise price, thereby achieving the joint profit maximizing result, but there is an incentive on the part of the members to increase individual profit. If this kind of situation occurs, it leads to break-up of the cartel. The difficulty with sustaining collusion is often demonstrated by a classic strategic game known as the prisoner's dilemma.

Game theory (Prisoner's dilemma)

The story is something like this. Two KGB officers spotted an orchestra conductor examining the score of Tchaikovsky's Violin Concerto. Thinking the notation was a secret code, the officers arrested the conductor as a spy. On the second day of interrogation, a KGB officer walked in and smugly proclaimed, "OK, you can start talking. We have caught Tchaikovsky".

More seriously, suppose the KGB has actually arrested someone named Tchaikovsky and the conductor separately. If either the conductor or Tchaikovsky falsely confesses while the other does not, the confessor earns the gratitude of the KGB and only one year in prison, but the other receives 25 years in prison. If both confess each will be sentenced to 10 years in prison; and if neither confesses each receives 3 years in prison. Now consider the outcome.

The conductor knows that if Tchaikovsky confesses, he gets either 25 years by holding out or 10 years by confessing. If Tchaikovsky holds out, the conductor gets either 3 years by holding out or only one year confessing. Either way, it is better for the conductor to confess. Tchaikovsky, in a separate cell, engages in the same sort of thinking and also decides to confess. The conductor and Tchaikovsky would have had three-years rather than 10-year jail sentences if they had not falsely confessed, but the scenario was such that, individually, false confession was rational. Pursuit of their own self interests made each worse off.

CONDUCTOR

		Cooperate	Confess
TCHIKOVSKY	Cooperate	3.3	25.1
	Confess	1.25	10,10

This situation is the standard prisoner's dilemma and is represented in the above matrix. This first payoff in each cell refers to Tchaikovsky's, and the second is the conductors. Examination of the payoffs shows that the joint profit maximizing strategy for both is (Cooperate-Cooperate).² The assumption in this game is that both the parties decided their strategies independently. Let us assume both parties are allowed to consult each other before the interrogation. Do you think cooperation will be achieved? It is unlikely since each of them will individually be concerned about the worst outcome that is 25 years in jail. Cooperation in this prisoner's dilemma becomes even more difficult, because it is a one shot game.

This scenario is easily transferred to the pricing decision of a company. Consider two companies setting prices. If both companies would only keep prices high, they will jointly maximize profits. If one company lowers price, it gains customers and it is thus in its interests to do so. Once one company has cheated and lowered price, the other company must follow suit. Both companies have lowered their profits by lowering price. Clearly, companies repeatedly interact with one another, unlike Tchaikovsky and the conductor. With repeated interaction, collusion can be sustained.

Robert Axelrod, a well-known political scientist, claims a "tit-for-tat" strategy is the best way to achieve co-operation. A tit-for-tat strategy always co-operates in the first period and then mimics the strategy of its rival in each subsequent period.

Axelrod likes the tit-for-tat strategy because it is nice, retaliatory, forgiving and clear. It is nice, because it starts by co-operating, retaliatory because it promptly punishes a defection, forgiving because once the rival returns to co-operation it is willing to restore co-operation, and finally its rules are very clear: precisely, an eye for an eye.

A fascinating example of tit-for-tat in action occurred during the trench warfare of the First World War. Front-line soldiers in the trenches often refrained from shooting to kill, provided the opposing soldiers did likewise. This restraint was often indirect violation of high command orders.

Price Leadership

Price leadership is an alternative cooperative method used to avoid tough competition. Under this method, usually one firm sets a price and the other firms follow. It is quite popular in industries like cigarette industry. Here any firm in the oligopolistic market can act as a price leader. The firm, which is highly efficient, and having low cost can be a price leader or the firm, which is dominant in the market acts as a leader. Whatever the case may be, the firm, which sets the price, is the price leader. We have two forms of price leadership-**Dominant price leadership** and **Barometric price leadership**.

² Remember the payoffs in the matrix are years in jail, thus the lesser the better.

In dominant price leadership, the largest firm in the industry sets the price. If the small firms do not conform to the large firm, then the price war may take place due to which the small firms may not be able to survive in the market. It is more or less like a monopoly market structure. This can be seen in the airlines industry in India where the dominant Airlines Firms sets prices and the others Subordinate Airlines Firms follow the price changes of dominant Airlines Firms.

Barometric price leadership is said to be the simpler of the two. This normally occurs in the market where there is no dominant firm. The firm having a good reputation in the market usually sets the price. This firm acts as a barometer and sets the price to maximize the profits. Here it is important to note that the firm in question does not have any power to force the other firms to follow its lead. The other firms will follow only as long as they feel that the firm in action is acting fairly. Though this method is quite ambiguous regarding price leadership, it is legally accepted. These two forms are an integral part of different types of cooperative oligopoly. Barometric price leadership has been seen in the automobile sector.

ILLUSTRATION

Reestablishing Price Discipline in the Steel industry

A US steel company S was the leader in setting prices in the steel industry. However, in 1962, a price increase announced by S company provoked so much criticism from customers and elected officials, that the firm became less willing to act as the price leader. As a result, the industry evolved from dominant firm to barometric price leadership. This new form involved one firm testing the waters by announcing a price change and then S. Steel company either confirming or rejecting the change by its reaction.

Later S company found that its market share was declining. The company responded by secretly cutting prices to large customers. This action was soon detected by Steel company B which cut its posted price of steel from \$113.50 to \$88.50 per ton. Within three weeks, all of the other major producers, S Steel included, matched Steel company B's new price.

The lower industry price was not profitable for the industry members. Consequently, U.S. Steel signaled desire to end the price war by posting a higher price. Steel company B waited nine days and responded with a slightly lower price than that of S. Steel. S. Steel was once again willing to play by industry rules.

Steel company B announced a price increase to \$125 per ton. All of the other major producers quickly followed suit, and industry discipline was restored. Note that the price of \$125 per ton was higher than the original price of \$113.50.

Source: Peterson and Lewis, 2002. Managerial Economics. Pearson Education Asia.

Activity 3

Suppose a firm is operating in a non-cooperative oligopolistic market structure. It produces 400 units of output per period and sells them at Rs. 5 each. At this stage its total revenue is Rs. 2,000. The firm now thinks of changing its price and increases it from Rs. 5 to Rs. 6. The rivals do not change the price and the sales dip from 400 units to 200 units.

Now the firm decides to decrease the price of the product from Rs. 5 to Rs. 4 and expects the rivals to match the price decrease so as not to lose sales. Now the sales increase marginally from 400 units to 450 units.

- a) Find the total revenue of the firm when its price increases from Rs. 5 to Rs.6.

.....

.....

.....

.....

.....

.....

.....

- b) Find the total revenue of the firm when the price decreases from Rs.5 to Rs.4.

.....

.....

.....

.....

.....

.....

.....

- c) Plot the changes along with the initial price and quantity sold according to the concept to price rigidity under oligopoly.

.....

.....

.....

.....

.....

.....

.....

12.6 SUMMARY

In this unit we have tried to explain the concept of pricing with special reference to monopolistic and oligopolistic competition. The effort has also been made to include the application part of the concept of product differentiation in monopolistic competition and the oligopolistic competition. Talking about monopolistic competition, we have seen that in monopolistic competition, the firm's economic profit is evaded off completely in the long run. In the short run monopolistic competition is quite similar to monopoly. We have discussed the oligopolistic competition in brief. The main characteristic of oligopolistic competition seems to be mutual interdependence and this factor decides the nature of oligopolistic competition.

We can summarize the whole unit by saying that the basis of differentiation between different types of competitions comprises of the number of sellers, the number of buyers, product differentiation, and barriers to entry. These factors decide the nature of competition in a particular market structure.

12.7 KEY WORDS

Marginal Revenue is the revenue obtained from the production and sale of one additional unit of output.

Marginal cost is the cost arising due to the production of one additional unit of output.

Economic profit is also known as the pure profit and is the residual left after all contractual costs have been met.

Price leadership a firm setting up the price at profit maximizing level and other firms following it.

Non-price competition is a form of competition used in Oligopolistic competition where price change by firms is not involved.

12.8 SELF-ASSESSMENT QUESTIONS

1. Distinguish between perfect competition and imperfect competition, giving examples.
2. Which of the following markets could be considered monopolistically competitive? Explain.
Cable Television
Ball pens (low priced)
Food joints
Automobiles
3. Take the case of a monopolistically competitive firm and describe the steps involved in attaining long-run equilibrium for the firm.

4. Explain whether the firms producing differentiated products are more likely to face price competition than the oligopolists producing homogeneous products.
5. Write short notes on:
Dominant price leadership
Barometric price leadership
6. Which of the following markets could be considered oligopolistically competitive? Explain.
Theaters
Automobiles
Aircrafts
Restaurants
Oil producing companies
Yarns
Newspapers
Garments
Cereals
Branded products like Photo film
7. Suppose production decisions of two members of OPEC, say Iran and Iraq are as follows. Each has just two production levels, either 2 or 4 million barrels of crude oil a day. Depending on their decisions, the total output on the world market will be 4, 6, or 8 million barrels. Suppose the price will be \$25, \$15, and \$10 per barrel, respectively. Extraction costs are \$2 per barrel in Iran and \$4 per barrel in Iraq.
 1. Represent the game in the form of a Prisoner's Dilemma.
 2. If Iran were to cheat successfully, what would be the daily increase in Iran's profits?
 3. If Iraq were to cheat successfully, what would be the daily increase in Iraq's profits?
 4. For which of the countries is the cost of cheating higher. Why?
 5. If it takes Iraq a month to detect Iran's cheating and respond, how many days will it take for the extra profits of Iran to be wiped out?
 6. What are some of the mechanisms you can think of that will entice co-operation from the two countries.

12.9 FURTHER READINGS

Mote, V. L., Paul, S., & Gupta, G. S. (2016). Managerial Economics: Concepts and Cases, Tata McGraw-Hill, New Delhi.

Maurice, S. C., Smithson, C. W., & Thomas, C. R. (2001). Managerial Economics: Applied Microeconomics for Decision Making. McGraw-Hill Publishing.

Dholakia, R., & Oza, A. N. (1996). *Microeconomics for Management Students*. Oxford University Press, Delhi.

Lewis, W. C., Jain, S. K., & Petersen, H. C. (2005). *Managerial Economics* (4th ed.). Pearson



UNIT 13 PRICING STRATEGIES

Objectives

After going through this unit, you should be able to:

- understand different pricing strategies adopted by firms;
- identify the relevance of these pricing strategies under different conditions;
- apply pricing decision store world situations.

Structure

- 13.1 Introduction
- 13.2 Concentration Ratios, Herfindahl Index & Contestable Market
- 13.3 Price Discrimination
- 13.4 Peak Load Pricing
- 13.5 Bundling
- 13.6 Two-Part Tariffs
- 13.7 Pricing of Joint Products
- 13.8 Transfer Pricing
- 13.9 Other Pricing Practices
- 13.10 Summary
- 13.11 Key Words
- 13.12 Self-Assessment Questions
- 13.13 Further Readings

13.1 INTRODUCTION

Not every customer is willing to pay the same price for the same product. So how is a seller to set prices to maximize business? The answer is the world of price discrimination. A good example of price discrimination is different prices charged for movie, concert tickets from students and adults. U Cab company is an American company which offers cab services across the world, even in India. Price discrimination is one of the major strategies they have been following for a while now. Pricing for a same distance journey is different for different customers depending on their geographical location. The question “How should a product be priced?” is of enormous importance to businesses, and most companies allocate substantial budgets to market research, both before launching a new product and, once launched, through the different stages of the product’s life cycle. Economists argue that the level of demand for a product at any price is the sum of what all individual

consumers in the market would be willing to purchase. This demand or willingness to pay, for any product is affected by three key factors:

- Individual consumers' preferences for the different characteristics of the product.
- The price of close substitutes to the product and the price of goods that must be used in conjunction with it.
- The level of each individual consumer's income.

This will apply to any product, be it cans of softdrinks, automobiles or laptops. This unit will examine the common pricing strategies adopted by firms including price discrimination.

13.2 CONCENTRATION RATIOS, HERFINDAHL INDEX AND CONTESTABLE MARKETS

The degree by which an industry is dominated by a few large firms is measured by **Concentration ratios**. These give the percentage of total industry sales of 4, 8, or 12 largest firms in the industry. An industry in which the four-firm concentration ratio is close to 100 is clearly oligopolistic, and industries where this ratio is higher than 50 or 60 percent are also likely to be oligopolistic. The four-firm concentration ratio for most manufacturing industries in the United States is between 20 and 80 percent.

Another method of estimating the degree of concentration in an industry is the **Herfindahl index (H)**. This is given by the sum of the squared values of the market shares of all the firms in the industry. The higher the Herfindahl index, the greater is the degree of concentration in the industry. For example, if there is only one firm in the industry so that its market share is 100%, $H=100^2=10,000$. If there are two firms in an industry, one with a 90 percent share of the market and the other with a 10 percent share, $H = 90^2 + 10^2 = 8,200$. If each firm had a 50 percent share of the market, $H = 50^2 + 50^2 = 5,000$. With four equal-sized firms in the industry, $H = 2,500$. With 100 equal-sized firms in the (perfectly competitive) industry, $H = 100$. This point to the advantage of the Herfindahl index over the concentration ratios discussed above. Specifically the Herfindahl index uses information on all the firms in the industry- not just the share of the market by the largest 4, 8, 12 firms in the market. Furthermore, by squaring the market share of each firm, the Herfindahl index appropriately gives a much large weight to larger than to smaller firms in the industry. The Herfindahl index has become of great practical importance since 1982 when the Justice Department in the US announced new guidelines for evaluating proposed mergers based on this index.

In fact, according to the theory of **Contestable markets** developed during the 1980s, even if an industry has a single firm (monopoly) or only a few firms (oligopoly), it would still operate as if it were perfectly competitive if entry is

“absolutely free” (i.e. if other firms can enter the industry and face exactly the same costs as existing firms) and if exit is “entirely costless” (i.e., if there are no sunk costs so that the firm can exit the industry without facing any loss of capital). An example of this might be an airline that establishes a service between two cities already served by other airlines if the new entrant faces the same costs as existing airlines and could subsequently leave the market by simply reassigning its planes to other routes without incurring any loss of capital. When entry is absolutely free and exit is entirely costless, the market is contestable. Firms will then operate as if they were perfectly competitive and sell at a price which only covers their average costs (so that they earn zero economic profit) even if there is only one firm or a few of them in the market.

13.3 PRICE DISCRIMINATION

In economic jargon, price discrimination is usually termed monopoly price discrimination. This label is appropriate because price discrimination cannot happen in a perfectly competitive industry in equilibrium. Monopoly power must be present in a market for price discrimination to exist. This seems a trivial point, when you understand, the definition of price discrimination; the practice of charging different prices to various consumers for a given product. In a competitive market, consumers would simply buy from the cheapest seller, and producers would sell to the highest bidders, and that would be that.

With monopoly power, however, the opportunity may exist for the firm to offer different terms (of which price is only one component) to different purchasers, thus dividing the market—a practice known as market segmentation. Price discrimination refers to the situation where a monopoly firm charges different prices for exactly the same product. The monopoly firm (a single seller in the market) can discriminate between different buyers by charging them different prices because it has the power to control price by changing its output. The buyers of its product have no choice but to buy from it as the product has no close substitutes.

There are three types of price discrimination – First Degree price discrimination, Second Degree price discrimination, and Third Degree price discrimination. **First degree price discrimination** refers to a situation where the monopolist charges a different price for different units of output according to the willingness to pay of the consumer. For example, a doctor who is the only super specialist in the town may charge different fee for conducting surgery from different patients based on their ability to pay. **Second degree price discrimination** refers to a situation where the monopolist charges different prices for different set of units of the same product. For example, the electricity charges per unit of the first 100 Kwh of power consumption may be different from the rate charged for the additional 100 Kwhs. Another example is railway passenger fares; the per kilometer

fare is higher for the first few kilometers, which declines as the distance increases. Thus the discrimination is based on volume of purchases. When the monopolist firm divides the market (for its product) into two or more markets (groups of buyers or segments) and charges different price in each market, it is known as **third degree price discrimination**. Airline tickets are a common example of this form of price discrimination. For example, lower rates are applicable to senior citizens than business travelers, electricity rates applicable to residential users are lower than those applied to commercial establishments and soon.

a) First Degree Price Discrimination

Monopolists engage in price discrimination when they can increase their profits by doing so. Even if sellers know the maximum amount that different customers are willing to pay, developing a pricing scheme that makes each customer pay that amount, a practice known as first degree price discrimination, can be difficult. Under first degree price discrimination, the full benefit from the trade between buyer and seller accrues to the seller. One strategy to achieve first degree price discrimination is to sell to the highest bidders through sealed bid auctions. The auction approach is best suited for situations where the volume of sales are low (usually due to scarcity of the product), where there are many potential buyers who are unable to co-operate among themselves and where buyers all have access to the same information about the product's characteristics. The auction approach would enable to seller to identify those buyers with the highest willingness to pay and would yield the highest possible revenues for the same production costs. This is a common strategy for the sale of very special types of products such as art objects, antique furniture or the rights to the mining and exploration of plots of land. It is not suitable for most bulk-produced products such as cans of soft drinks or computers. Perfect or first-degree price discrimination can occur when a firm knows the maximum price the individual is willing to pay for each successive unit. The firm could then charge that highest price for each successive unit and capture the entire consumer surplus. Remember that all forms of price discrimination involve some monopoly power, but perfect price discrimination involves a degree of monopoly power rarely found in the real world.

b) Second Degree Price Discrimination

Where the auction approach is not feasible, the company must do its best to approximate the first degree outcome using its pricing structure. This is based on the notion that an individual consumer derives diminishing satisfaction from each successive unit of any product consumed.

This form of price discrimination, which is based on the volume of consumer purchases, is very common and is known as *second degree price discrimination*. Other forms of second degree price discrimination include two-tier tariffs, i.e. prices where the consumer must pay a flat fee for access

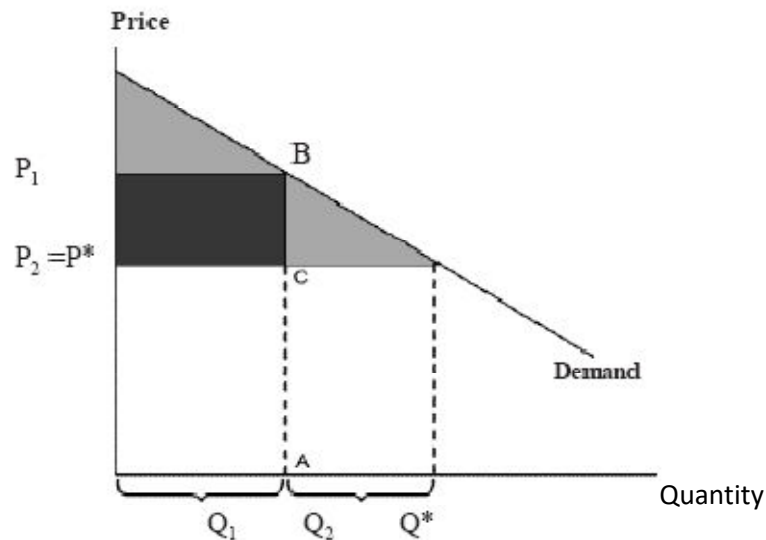
and then a separate fee (which may be zero) for usage. This is typical of many clubs, amusement parks and transport facilities offering monthly or annual passes.

The idea in the case of travel pass, for example, is that the traveller who travels infrequently pays on average, a higher price per trip because the fixed access cost is spread over fewer trips. On the other hand, the high volume user spreads this fixed cost over so many trips that he or she may actually sit next to the infrequent traveller, consume the exact same services (meals, fuel and so on), but end up paying a lower average price for any given trip.

Second-degree price discrimination is also referred to as **multipart pricing**. It is a block, or step, type of pricing, in which the first set of units is sold at one price, a second set at a lower price, a third set at a still lower price, and so on. Note that this is different from a quantity discount in which the lower (discounted) price applies to all units purchased. In second-degree price discrimination, the lower price applies only to units purchased in that block. The buyer must have already paid the higher price for the earlier units. Some familiar examples should make this clear:

1. Many retailers adopt second degree price discrimination for encouraging customers to buy in bulk. E.g., when a customer buys 1 shirt s/he will be charged full MRP for that, say rupees 2000. But if s/he buys one more shirt then s/he will be given discount of rupees 1000 on that. So, the first shirt will be sold at rupees 2000 and second shirt will be sold at rupees 1000.
2. Take example of pricing of an American fast food company which sells burger or fries. When you buy individual burger or fries, you are charged more as compared to when you buy burger fries and soft drink together as a meal.

Now, let's look at second-degree price discrimination in a more formal graphic model. In figure 13.1, the seller faces the demand curve (D) of one typical consumer. Although the cost function is not shown in the figure, assume that marginal revenue and marginal cost intersect and lead to an optimal price of P^* . The consumer would choose to buy the quantity Q^* at this price. The shaded area of the figure represents the consumer's surplus. It may be, however, that the firm uses multi part pricing to capture a portion of this surplus. Suppose that the firm sets a price of P_1 for the first Q_1 units purchased and that additional units sell for P_2 (a two-stage pricing scheme). The consumer buys Q_1 units at price P_1 and Q_2 units at price P_2 . That portion of the consumer surplus labeled P_1BCP_2 is now captured by the firm rather than by the consumer. This still leaves a rather large portion of the consumer surplus still in the consumer's hands. The firm's management would prefer to capture it all, and could do so by using more parts in a multipart pricing strategy. However, to do so, management needs to know a great deal about the consumer's demand.

Figure 13.1: Second-Degree Price Discrimination

In this example of second-degree price discrimination, or multi part pricing, the first block of units (Q_1 units) is sold at the price P_1 , and the second block (Q_2 units) is sold at the price P_2 . This allows the seller to capture that part of the consumer's surplus represented by the area P_1BCP_2 .

c) Third Degree Price Discrimination

Pricing based on what type of consumer is doing the purchasing rather than the volume of purchase is an approach known as third degree price discrimination. This is very common in the sales of air and rail travel, movie tickets and other products where consumers can be segmented into different groups, who are likely to differ greatly in their willingness to pay based on certain easily identifiable attributes.

Thus, third-degree price discrimination, or market segmentation, requires that the seller be able to (1) segment, or separate, the market so that goods sold in one market cannot be resold by the buyers in another; and (2) identify distinct demand curves with different price elasticities for each market segment.

Students are one of the main beneficiaries of third degree price discriminations schemes, since their demand is more sensitive than the population at large. Other often identified groups include senior citizens and the young, both of whom also tend to be more price sensitive, and business purchasers, who are often less price sensitive and may be willing to pay a lot for small quality improvements. Suppose, for example, there are only two types of travellers; students and business men.

Students pay for their travel out of their own pockets, while businessmen charge their travel to their employers who in turn deduct these expenses from their taxable income. Since a typical student is likely to be willing to pay less for a travel ticket, all else being equal, than a typical businessmen, it makes

sense for the company selling travel services to price higher to the businessman and lower to the tourist to get the largest possible volume of business out of each customer group.

Pricing schemes can be quite complex and may combine elements of second- and third-degree price discrimination: for example, discounted travel passes for students and pensioners. In any case, the main danger to the seller is that customers have an incentive to get together and trade among themselves to benefit from existing price differentials.

Thus, a student may try to purchase a ticket s/he does not plan to use for the express purpose of selling it to a business traveler and sharing the difference between the prices. Or, a holder of a travel pass may offer the pass to a friend to use, enabling the friend to benefit from the high volume of the holder's travel. If this were allowed to happen, the seller would lose the business of the high-price paying customer and would be better off offering a single profit-maximizing price.

The seller engaging in price discrimination must therefore take measures such as passport checks at the departure gate and photos on rail passes to make sure consumers are not able to engage in *arbitrage*, i.e. profit from their access to a lower price by selling to someone to whom such access is precluded.

The other danger the price discriminating seller faces is that a rival firm may enter with a single price that undercuts the incumbent's higher price. Then the rival will draw away the most profitable market segments and the original company will only be left with the low-margin discount buyers.

That is why price discrimination is only possible in imperfectly competitive markets, where direct competition by rivals is made difficult by entry barriers such as established brand names (laptops), differentiated products (magazines), scale economies in production (air and rail travel), technology patents (pharmaceuticals) for where access to a key input is limited (fine art).

Activity 1

1. Necessary conditions for price discrimination include all, but the following.
 - a) The firm must have some control over price.
 - b) The firm's markets must be separable.
 - c) The firm must have declining long run average costs.
 - d) The elasticity of demand must vary among markets.
2. The train fare charged by Delhi Metro from Dwarka to Saket is Rs. 45 during the morning rush hour from 9:00 am to 11:00 am, but drops to Rs. 40 after 11:00 am. This is because the demand for train rides from Dwarka to Saket is:

- a) Elastic in the rush hour, but inelastic later in the day.
 - b) Unit elastic at all times of the day.
 - c) Inelastic in the rush hour, but elastic later in the day.
 - d) Unit elastic in the rush hour, but inelastic later in the day.
3. Price discrimination is possible between markets. Discuss, with help of an example

.....

.....

.....

.....

.....

13.4 PEAK LOAD PRICING

Peak load pricing is a type of third-degree price discrimination in which the discrimination base is temporal. We single out this particular form of price discrimination in part because of its widespread use. But remember that all forms of third-degree price discrimination, including peak load pricing, involve a seller attempting to capitalize on the fact that buyers' demand elasticities vary. In the case of peak load pricing, customer demand elasticities vary with time.

Very few, if any, business economic activities are characterized by an absolutely constant demand during all seasons of the year and at all times of day. For many, the variations, or fluctuations, are not large enough to be of concern; but for some activities, fluctuations in demand are significant. These variations are sometimes relatively stable and predictable. Hotel booking prices for vacation spots (say hill stations in India) are good example of peak load pricing. During peak season (summers), hotels booking charges are highest and reduced during off season. Travelers are encouraged to visit hill station during slack period (off season) and those who visit during peak season pay relatively higher prices. Whenever price discrimination is based on time differentials, the object of the selling firm is to charge a higher price for the product during the more inelastic period and a lower price during the more elastic interval.

Activity 2

1. Calculus can be used to accomplish price discrimination provided that the firm knows its TC functions and the demand functions for the markets it is selling to. Consider two markets X and Y. The total revenue functions are TR_X and $TR_Y = P_X Q_X$ and $P_Y Q_Y$ respectively. The demand curves for the two markets are :

$$P_x = 2 - Q_x$$

$$P_y = 3 - 2Q_y$$

While the producers marginal cost is given by

$$TC = 100 + 1.5Q$$

$$\text{Where } Q = Q_x + Q_y$$

Calculate the profit maximizing output and price with and without price discrimination.

Show that profit is larger when the firm practices price discrimination.

.....

2 Peak load pricing is a type of third-degree price discrimination. Recall that all forms of third-degree price discrimination, including peak load pricing, involve a seller attempting to capitalize on the fact that buyers' demand elasticities vary. In the case of peak load pricing, customer demand elasticities vary

- a) Across categories of consumers
- b) With time.
- c) Across different regions or countries

13.5 BUNDLING

You must have come across campaigns of the following kind. "Buy one, get the second at half-price". A camera is sold in a box with a free tripod; a hotel room often comes with complimentary breakfast. These are examples of Bundling.

Bundling is the practice of selling two or more separate products together for a single price i.e. bundling takes place when goods or services which could be sold separately are sold as a package. A codification of bundling practices and definitions of selling strategies is:

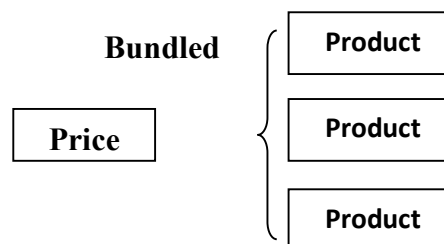
Pure bundling: products are sold only as bundles;

Mixed-bundling: products are sold both separately and as a bundle; and

Tying: The purchase of the main product (tying product) requires the purchase of another product (tied product) which is generally an additional complementary product.

This is not an exhaustive list but covers the most frequently encountered cases. Pure bundling involves selling two products only as a package and not separately.

PURE BUNDLING

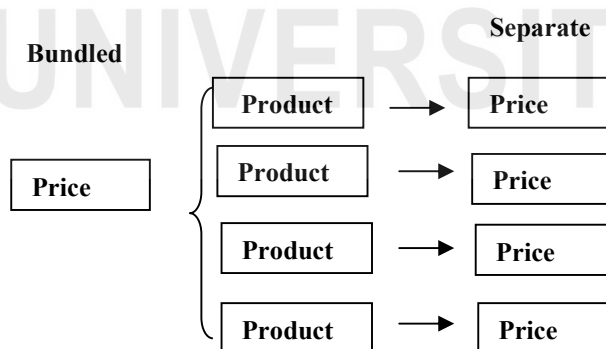


For example, Microsoft Office 365 is available as bundle of different applications such as word, excel, access, PowerPoint, outlook, OneNote etc. You cannot buy these applications separately, you have to buy this in bundle only. Microsoft has followed this practice because some applications are more in demand as compared to others.

Mixed Bundling involves selling products separately as well as a bundle.

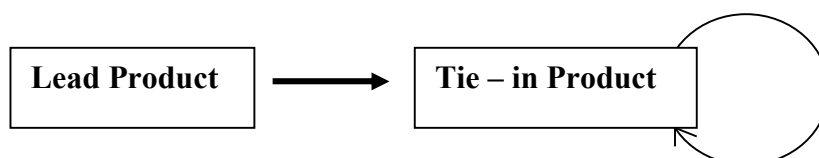
An American fast food company's are examples of Mixed Bundling. The Times of India and The Economic Times can be purchased together for weekdays for a price much less than if purchased separately. This is also an example of mixed bundling. In most cases mixed bundling provides price savings for consumers.

MIXED BUNDLING



Tying involves purchases of the main product (tying product) along with purchase of another product (tied product) which is generally an additional complementary product.

TYING



Customers who leased a photo Copying Machine had to buy Photo copy Paper. Another case of tying was that by Kodak in which Kodak held a monopoly in the market for Kodak Copier Parts. Kodak engaged in tying when it refused to sell its parts to consumers or independent service providers except in connection with a Kodak Service Contract. Today when you buy a Mach3 razor, you must buy the tied product i.e. the cartridge that fits into the Mach3 razor.

Financial bundling has become widespread. “Manufacturing is becoming the loss-leader of the profit chain for many companies.” In other words, give away the product; make money on the lending that is bundled with it. In India too, a number of automobile companies are providing finance and bundling the automobile with financing.

Bundling can be good for consumers. It can reduce “search costs” (the bundled goods are in the same place), as well as the producer’s distribution costs. There are lower “transaction costs” (because a single purchase is cheaper to carry out than multiple ones). And the producer may be a more efficient bundler than the customer: few of us choose, after all, to buy the individual parts of a computer to assemble them ourselves.

In perfectly competitive markets, bundling should happen only if it is more efficient than selling the products separately. Where there is less than perfect competition—that is, most markets-economic models suggest that bundling sometimes benefits consumers and sometimes producers. When firms have a measure of market power, they can engage in price discrimination, charging different prices to different customers. Bundling can play a part in price discrimination, as different bundles of goods and prices may appeal to different customers.

In a celebrated case that caught much media attention, Microsoft was accused of anti-competitive conduct in ‘bundling’ Internet Explorer and Windows as a pure bundle. Microsoft claimed they are not a bundle at all, rather a single product incapable of being broken into parts. It is of course difficult to settle such arguments. But the interesting aspect is that the company does not consider its product (Windows and Internet Explorer) as being capable of being broken into parts.

Activity 3

1) What is bundling? Give examples. Do you think this is anti-consumer?

.....

.....

.....

.....

- 2) Give examples of tying from the Indian market.

.....

.....

.....

.....

13.6 TWO-PART TARIFFS

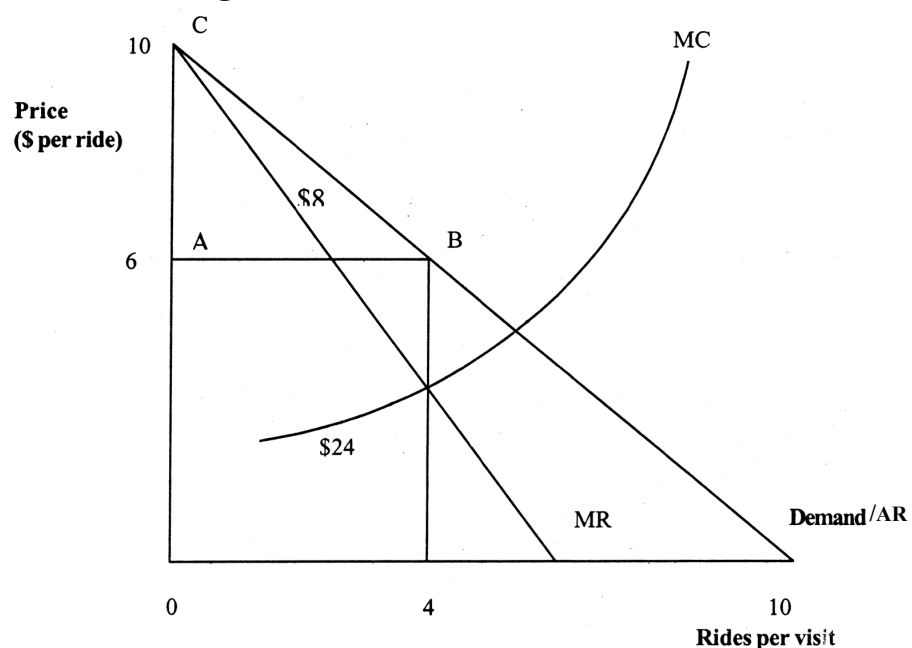
One of their techniques requires buyers to pay a fee for the right to purchase their product and then to pay a regular price per unit of the product. For example, mobile telecom operators A, V etc. used to charge a monthly fixed charge and then additional charges based on usage for their post-paid plans.

The fee for privilege of service plus prices for services consumed is called a two-part tariff. Theme parks such as D W usually employ such a pricing scheme to increase their profits. To see how the scheme works, suppose you operate a theme park and have a local monopoly. Figure 13.3 shows the demand for rides at your theme park by any given tourist, along with the marginal revenue and marginal cost of the rides. If you charge a single monopoly price, your rides will be priced at \$6 each and each tourist will consume four rides per visit, spending \$24.

Now let's see if a bit more can be extracted from each tourist. Given the demand curve drawn, each tourist would be willing to pay more than \$24 to enter your theme park and take four rides. If you know the demand curve for rides, you know that the typical tourist is enjoying a consumer surplus of \$8, corresponding to the area of triangle ABC in the graph (area $ABC = \frac{1}{2} \times 4 \times 4$). Therefore, if you charge an entry fee of \$8 in addition to \$6 per ride, you can add \$8 per tourist to your profit.

Given the demand curve of a typical tourist, you can add still more to your revenue from each tourist if you simply eliminate the price per ride and just charge an

Figure 13.3: Demand for rides



admission fee equal to consumer surplus at zero price per ride. For example, if the price per ride were zero, a tourist would go on 10 rides per visit and you would get revenue of \$50 per tourist $-0.5 (\$10)(10)$ —instead of the \$32 you would get from the two-part pricing scheme. But be careful. With more rides your marginal costs will increase, and thus your profit might not increase. Also, if you extract the entire consumer surplus with a single entry fee, you increase the tourists' cost per visit, so the total number of admissions will fall.

A two-part tariff is often a good way to increase profit by extracting some, but not all, of the consumer surplus from a monopolist's clients. A two-part tariff is often a good way to increase profit by extracting some, but not all, of the consumer surplus from a monopolist's clients. Monopolists usually experiment with various two-part tariff pricing schemes before hitting on the one that gives them maximum profit.

When theme Parks Amusement rides opened, it was the first amusement park of its kind in Mumbai and so had no precedent to go by. As the objective was to sell the concept to as many people as possible, it avoided charging a composite fee, for a stiff entrance fee would keep families away. Instead, it selected what seemed the most sensible approach: pay-as-you-go. It charged an entrance fee of ₹ 5 for children and ₹ 10 for adults. And the individual rides were priced between ₹ 2 and ₹ 15. Later, however, theme Parks Amusement rides jettisoned the split pricing strategy and switched over to composite pricing. Under the new tariff structure children were charged a fee of ₹ 80, while adults had to pay ₹ 100. There were no charges levied on the rides. What prompted the switch? According to their vice-president they found difficult to implement the pay-as-you-go strategy because of logistical problems.

13.7 PRICING OF JOINT PRODUCTS

Products can be related in production as well as demand. One type of production inter dependency exists when goods are jointly produced in fixed proportions. The process of producing leather shoes and handbags in a leather factory is a good example of fixed proportions in production. Each unit provides a certain amount of leather shoes and handbags. There is little that the leather factory can do to alter the proportions of the two products.

When goods are produced in fixed proportions, they should be thought of as a "product package." Because there is no way to produce one part of this package without also producing the other part, there is no conceptual basis for allocating total production costs between the two goods. These costs have meaning only in terms of the product package.

Calculating the Profit-Maximizing Prices for Joint Products

Assume a retailer sells leather shoes and handbags. The two goods are assumed to be jointly produced in fixed proportions. The marginal cost equation for the product package is given by

$$MC = 30 + 5Q$$

The demand and marginal revenue equations for the two products are

Shoes	Handbags
$P = 60 - 1Q$	$P = 80 - 2Q$
$MR = 60 - 2Q$	$MR = 80 - 4Q$

What prices should be charged for shoes and handbags? How many units for the product package should be produced? Summing the two marginal revenue (MR_T) equations gives

$$MR_T = 140 - 6Q$$

The optimal quantity is determined by equating MR_T and MC and solving for Q . Thus

$$140 - 6Q = 30 + 5Q$$

and, hence, $Q = 10$

Substituting $Q = 10$ into the demand curves yields a price of \$50 for shoes and \$60 for handbags. However, before concluding that these prices maximize profits, the marginal revenue at this output rate should be computed for each product to assure that neither is negative. Substituting $Q = 10$ into the two marginal revenue equations gives 40 for each good. Because both marginal revenues are positive, the prices just given maximize profits. If marginal revenue for either product is negative, the quantity sold of that product should be reduced to the point where marginal revenue equals zero.

13.8 Transfer Pricing

In today's world, a lot of companies have divided their operations into several divisions. **Transfer pricing is setting the price for goods that are sold between these related legal entities (branches, subsidiary etc.).** It means the price of goods, which is paid by one unit of an organization to another unit. E.g., Branch A of a company sells goods to Branch B, the price paid by B is the transfer price. Transactions covered under transfer pricing are Sale of finished goods, Purchase of raw material, IT Enabled services, Purchase of fixed assets, Sale or purchase of machinery, Sale or purchase of Intangibles, Reimbursement of expenses paid/received, Support services, Software Development services, Technical Service fees, Management fees, Royalty fees, Corporate Guarantee fees, Loan received or paid etc.

- Profits of different units can be ascertained separately and this helps in separate performance evaluation of each unit of an organization.
- Transfer pricing also impact resource allocation among different units of an organization.

Organization for Economic Cooperation and Development (OECD) is responsible for governing transfer pricing for multinational companies. Through transfer pricing, multinationals try to take tax advantage and that is the reason OECD has set some guidelines. It has specified a particular principle named **Arm's Length Principle (ALP)** for these type of dealings between related parties. ALP states that prices should be set by assuming each party as independent, not dependent or related parties. In simple words, it means transactions should be executed at fair market price. In addition to this principle, guidelines also specify different methods of transfer pricing. These methods are divided into two main heads: **Traditional Transaction Methods and Transactional Profit Methods.**

Traditional Transaction Methods

- **Comparable Uncontrolled Price (CUP) method**
Under these methods, comparison is done between the terms and conditions of uncontrolled transactions (with unrelated parties) and controlled transactions (related parties). This requires high standard comparable data to ensure that transaction has been done on the basis of ALP.
- **Resale price method**
In this method, selling price of a product is used which is known as resale price. Margin amount and different costs (determined by comparing to unrelated transactions) are deducted to determine resale price.
- **Cost Plus method**
Under this, a markup (profit) amount is added in the total costs incurred by the supplier unit of the organization.

Transactional Profit Methods

- **The Comparable Profits Method**
Under this method, basis of the transfer price is net profit of controlled transaction.
- **The Profit Split Method**
This pricing method is used where companies are engaged in interconnected transactions and profits are separated by examining how unrelated parties would split profits in similar transactions.

13.9 Other Pricing Practices

Prestige Pricing

Prestige pricing is a type of pricing strategy, where a product is placed at a higher price, because consumers associate higher price with higher quality and a prestige is attached to the ownership of the product. This is also known as Image Pricing. NIKE is one of the best examples of prestige pricing.

Price Skimming

It is a strategy where a company charges high prices from consumers at first and earn substantial profits, later reduces the prices gradually to attract other customers who are price sensitive. APPLE is the best example of this type of pricing strategy.

Penetration Pricing

Companies charge relatively lower prices for their product to attract customers and increase market share. This pricing strategy is to lure customers away from competitor's products.

Psychological Pricing

Psychological pricing involves pricing that has impact on the psychology of the consumers. E.g., Companies set prices such as ₹ 499 etc. which consumers perceive as lower prices than they actually are.

13.10 SUMMARY

For a firm to be able and willing to engage in price discrimination, the buyers of the firm's product must fall into classes with considerable differences among classes in the price elasticity of demand for the product, and it must be possible to identify and segregate these classes at moderate cost. Also, buyers must be unable to transfer the product easily from one class to another, since otherwise persons could make money by buying the product from the low-price classes and selling it to the high-price classes, thus making it difficult to maintain the price differentials among classes. The differences among classes of buyers in the price elasticity of demand may be due to differences among classes in income, level, tastes, or the availability of substitutes.

13.11 KEYWORDS

First degree price discrimination refers to a situation where the monopolist charges a different price for different units of output according to the willingness to pay of the consumer.

Second degree price discrimination refers to a situation where the monopolist charges different prices for different set of units of the same product.

Third degree price discrimination is when the monopolist firm divides the market (for its product) into two or more markets (groups of buyers or segments) and charges different price in each market.

Peak load pricing is a type of third-degree price discrimination in which the discrimination base is temporal.

Bundling is the practice of selling two or more separate products together for a single price i.e. bundling takes place when goods or services which could be sold separately are sold as a package.

13.12 SELF-ASSESSMENT QUESTIONS

1. Assume a company produces a product that currently sells for ₹ 160. The unit costs for producing the product are—

Materials	₹ 51
Direct labour	₹ 32
Overhead	₹ 40
Sales expense	₹ 21
	₹ 144

These unit costs are based on sales of 100,000 units per year. Capacity is generally accepted to be 150,000 units per year. A foreign retail chain has contacted the company with an offer to purchase 60,000 units on a short-term basis during the next year at a price of 130 each. Sales of these units in the foreign market would not have any effect on the company's domestic market. Should the offer be accepted? Explain why or why not. Identify any assumptions you make in answering the question.

2. Which are the various methods of price discrimination identified in this unit? Explain with examples.
3. Why are auctions not used to extract consumer's surplus for most products sold? Under what conditions and for which goods are auctions useful to price the product being sold? Substantiate with a real world example.
4. Choose any product or service for which price discrimination exists in India. Identify the different categories of consumers and tabulate the corresponding prices for the chosen product or service. Comment on this pricing policy.
5. How many options does an amusement park have when it comes to the pricing decision?
6. Basically, there are four options open to the manager:
 - i) Keep the entrance free, but charge the visitor for every ride separately.

- ii) Charge a lump sum entrance fee and allow the visitor unlimited fee rides.
- iii) Same as the second option but limit the visitor to just one ride.
- iv) Charge an entrance fee and also get the consumer to pay a cumulative fee for the rides, the number of rides to be availed of being decided by the visitor.

The second, third and fourth options are based on the principle of a two-part tariff- the consumer is charged for entry and then for the rides (note that tariff for rides may be zero) . Traditional economic theory is able to conclusively prove that a two-part tariff is the best way to maximise revenue and hence the profits of the park owner, if two conditions simultaneously hold. First, if the seller is a monopolist. Second, the benefit of enjoying the good cannot be transferred. In case of an amusement park both these conditions hold. The rationale for a two-part tariff stems from the concept of consumer surplus. Consider a person who places the value for visiting the park and enjoying the rides at ₹ 100. Now if the rides are priced such that it costs him ₹ 90, then his (consumer) surplus is ₹ 10. He will opt for the product if the price is less than or equal to his perceived value.

- a) For the park owner, the trick is to extract a portion of this surplus. How?
- b) What are the variables you consider as important in the pricing decision?

13.13 FURTHER READINGS

Mote, V. L., Paul, S., & Gupta, G. S. (2016). *Managerial Economics: Concepts and Cases*, Tata McGraw-Hill, New Delhi.

Maurice, S. C., Smithson, C. W., & Thomas, C. R. (2001). *Managerial Economics: Applied Microeconomics for Decision Making*. McGraw-Hill Publishing.

Lewis, W. C., Jain, S. K., & Petersen, H. C. (2005). *Managerial Economics* (4th ed.). Pearson.

Keat, P. J., Young, P. KY., & Banerjee, S. (2009). *Managerial Economics: Economic Tool for Today's Decision Makers* (5th ed.). Pearson.